

Multivariate Geometric Expectiles

KLAUS HERRMANN*

MARIUS HOFERT[†]

MÉLINA MAILHOT[‡]

April 6, 2017

Abstract

A generalization of expectiles for d -dimensional multivariate distribution functions is introduced. The resulting geometric expectiles are unique solutions to a convex risk minimization problem and are given by d -dimensional vectors. They are well behaved under common data transformations and the corresponding sample version is shown to be a consistent estimator. We exemplify their usage as risk measures in a number of multivariate settings, highlighting the influence of varying margins and dependence structures.

Keywords: expectile, geometric quantile, elicibility, dependence, minimizing expected loss, multivariate risk measure

1 Introduction

A fundamental task in risk management and applied actuarial science is to quantify the risks associated with a given position. Prime examples of risky positions are portfolio holdings or (re-)insurance contracts. Quantifying risks is not only necessary for the internal decision making process of financial institutions, insurance companies or individual investors, but also mandatory from a regulatory perspective. For example, the regulatory frameworks for banks (OSFI, AMF, Basel II, 2.5, III) and insurance companies (CIA and, in Europe, Solvency II, Swiss Solvency Test) require not only internal risk modeling, but also specifically demand that businesses quantify and report risks in a specific way, using risk measures. This task is intrinsically multivariate in nature as one of the American Property and Casualty Minimum Capital Target Advisory Committee key principle is that ‘Risks should be aggregated. No diversification between risk categories is permitted until evidence confirms diversification will hold in a stress situation’ (see OSFI [2010]). The Office of the Superintendent of Financial Institutions of Canada states : ‘Gross, ceded and net provisions for claims liabilities must

*Department of Mathematics and Statistics, Concordia University, 1400 de Maisonneuve Blvd. West, Montréal (Québec) Canada H3G 1M8; e-mail: klaus.herrmann@concordia.ca

[†]Department of Statistics and Actuarial Science, University of Waterloo, 200 University Avenue West, Waterloo (Ontario), Canada N2L 3G1; e-mail: marius.hofert@uwaterloo.ca

[‡]Department of Mathematics and Statistics, Concordia University, 1400 de Maisonneuve Blvd. West, Montréal (Québec) Canada H3G 1M8; e-mail: melina.mailhot@concordia.ca.

be provided by actuarial lines of business’ (see Office of the Superintendent of Financial Institutions [2014]).

Until recently, regulatory economic capital has been calculated based on univariate risk measures. In this case, i.e., when considering risks separately, the theory of risk measures is well established, see, e.g., McNeil et al. [2015] Chapter 2 for an overview. The two most popular risk measures in this setting are value-at-risk (VaR) and tail-value-at-risk (TVaR; sometimes also referred to as conditional-tail-expectation or expected shortfall).

However, capital allocation has to be reinvestigated when dealing with a portfolio where pooling is not an option, or when it is more appropriate to secure capital simultaneously for multiple business activities. While the univariate framework is well understood it neglects interdependence and co-movements of different risks. In a real world scenario markets and assets are interconnected or prone to systemic risk. The same holds true when considering insurance contracts where dependence can play an important role. It can thus be beneficial to consider risks in a joint framework rather than treating them as isolated entities. Many problems from this consideration have been pointed out in the last decade, notably by Bank of Canada (see Gauthier et al. [2010]). To this end, a general theory for multivariate risk measures which specifically take the underlying dependence structure into account has recently emerged.

Based on multivariate risk measures, the trade-off between two stock indices has been studied by Cherubini et al. [2004] using bivariate inverse quantiles. Losses and adjustment loss allocation expenses (ALAE) have been studied by Di Bernardino et al. [2013], using multivariate value-at-risk and tail-value-at-risk. Guégan and Hassani [2011] allocate risk capital based on bivariate quantiles, where operational risk and other related risks are considered as separate dependent classes. Most of the techniques use an acceptance set, as presented in Jouini et al. [2004], and calculate a metric for each risk class, considering the dependence between those classes. Balbás et al. [2011] present several properties from a general representation of multivariate risk functions.

From an actuarial perspective, multivariate risk measures generalizing VaR are treated in Embrechts and Puccetti [2006], Cossette et al. [2012], Cousin and Di Bernardino [2013] and Torres et al. [2015]. Multivariate versions of TVaR have been defined in Cousin and Di Bernardino [2014] and Cossette et al. [2015]. Maume-Deschamps et al. [2017] also introduce a multivariate extension of expectiles. However, this approach differs from ours in the sense that it is non-geometrical. Likewise, the statistical community has generalized the notion of quantiles, i.e., VaR, to higher dimensions, e.g., via the notion of statistical depth, see, e.g., Mosler [2013] for an overview, and optimization-based definitions as in Abdous and Theodorescu [1992], Chaudhuri [1996] or Chakraborty [2001]. Although the two approaches set out from different starting points, interconnections are possible in some cases, see for example Hallin and Paindaveine [2010]. A thorough overview of different approaches can be found in Serfling [2002], while a connection between half-space depth and stress testing risk factors is established in McNeil and Smith [2012].

Our work is motivated by the fact that despite its good properties and popularity, the tail-value-at-risk is not elicitable in the univariate case, see Gneiting [2011]. Elicitability is a property that has been investigated in Osband [1985] in order to score the estimation of risks. Therefore, using the same criterion to do forecasting-based model selection and risk assessment is not possible when using TVaR, or, as shown in Ziegel [2014], any other

spectral risk measure other than the expectation. While univariate quantiles are elicitable, and can thus be utilized in forecasting-based model selection, they, however, do not adhere to the broadly accepted framework of coherence, see Artzner et al. [1999], which establishes preferable properties of risk measures in an axiomatic fashion. This is a serious drawback in actuarial applications.

As shown in Ziegel [2014], the only elicitable, law-invariant and coherent risk measures are expectiles, introduced by Newey and Powell [1987]. Expectiles generalize the mean for a given probability distribution in much the same way as quantiles generalize the median. Furthermore, they have a natural interpretation when considering the gain-loss ratio connected to a given position, i.e., the ratio of the expected gains over the expected losses, which is a popular performance measure in portfolio management, see Bellini and Di Bernardino [2017]. The amount of money that needs to be added to a given position in order to achieve a pre-specified, and in practical applications sufficiently high, gain-loss ratio is given by an expectile. In the univariate case, expectiles therefore combine favourable properties of risk measures and constitute an important addition to the well established VaR and TVaR.

Considering both, the need for multivariate risk measures and the favourable properties of univariate expectiles, the main target of the present study is therefore to define a multivariate version of expectiles and to study its properties.

The paper is structured as follows. We briefly summarize univariate quantiles and expectiles in Section 2, while the main ideas behind the multivariate framework are introduced in Section 3. Specifically, Section 3.1 reviews geometric quantiles, while Section 3.2 defines geometric expectiles as the main contribution of our paper. We discuss population and asymptotic properties of the newly introduced statistical functional in Section 4, while examples are discussed in Section 5. Finally Section 6 concludes. Selected plots can be reproduced with the latest version of the R package `qrmtools`; see the vignette `geometric_risk_measures`.

2 Univariate Quantiles and Expectiles

It is a standard approach in statistics to express population characteristics in terms of minimizing the expected loss of a random variable under a given loss function. Considering the absolute value $|\cdot|$, the median solves $\text{med } X = \operatorname{argmin}_{c \in \mathbb{R}} \mathbb{E}[|X - c|]$, while the mean is obtained when considering the square loss $\mathbb{E}[X] = \operatorname{argmin}_{c \in \mathbb{R}} \mathbb{E}[(X - c)^2]$. In case of the absolute value loss function it is readily observed that, using an asymmetric generalization of $|\cdot|$, quantiles other than the median can be obtained. For $\alpha \in (0, 1)$ we define the *check loss* as

$$\rho_\alpha: \mathbb{R} \rightarrow [0, \infty), \quad t \mapsto |\alpha - \mathbf{1}_{(-\infty, 0]}(t)| |t|, \quad (1)$$

where we see that the case $\alpha = 0.5$ is directly related to the usual absolute value. Similar to the median this leads to $F^{-1}(\alpha) = \operatorname{argmin}_{c \in \mathbb{R}} \mathbb{E}[\rho_\alpha(X - c)]$. In Koenker and Bassett [1978] this observation is the starting point to introduce the quantile regression framework. As an alternative Newey and Powell [1987], introduced an asymmetric version of the square loss along the same lines. To this end we set

$$\lambda_\alpha: \mathbb{R} \rightarrow [0, \infty), \quad t \mapsto |\alpha - \mathbf{1}_{(-\infty, 0]}(t)| |t|^2, \quad (2)$$

where again $\alpha \in (0, 1)$. The minimizers $E(\alpha) = \operatorname{argmin}_{c \in \mathbb{R}} \mathbb{E}[\lambda_\alpha(X - c)]$ are called expectiles, analogously to quantiles minimizing the check loss. Again the case $\alpha = 0.5$ reduces to the well known motivating example, i.e., $E(0.5) = \mathbb{E}[X]$. In Figure 1 the loss functions associated to quantiles and expectiles are shown for $\alpha \in \{0.5, 0.25, 0.9\}$. As clearly visible from Figure 1,

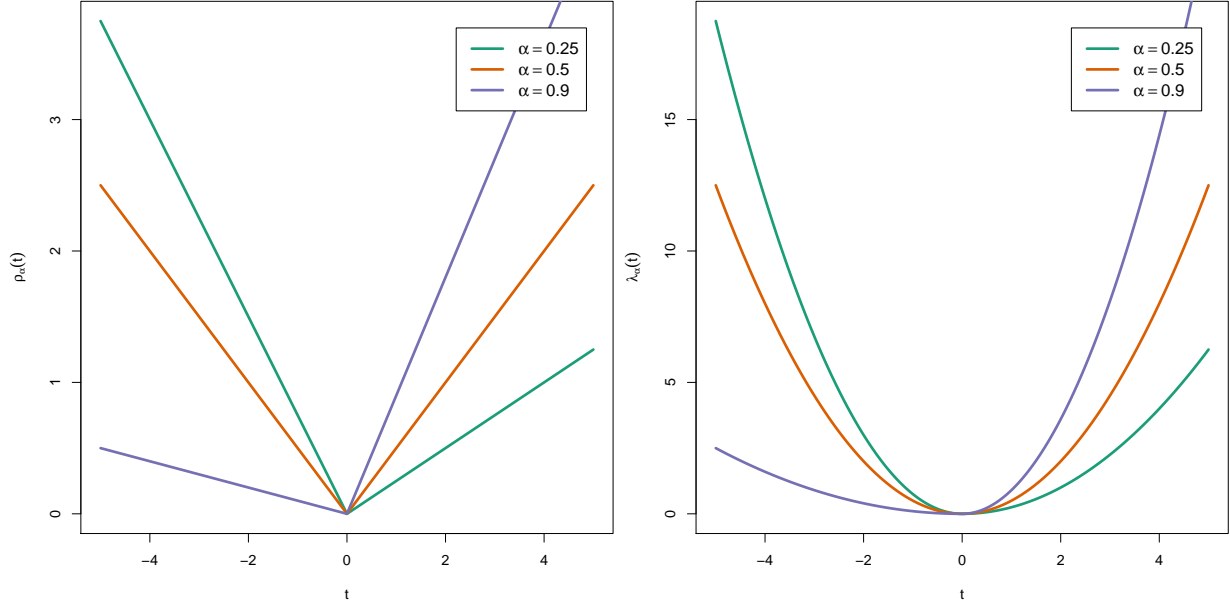


Figure 1: Loss functions ρ_α (left) and λ_α (right) for $\alpha \in \{0.25, 0.5, 0.9\}$.

the generalized loss functions are indeed asymmetric versions of their symmetric $\alpha = 0.5$ counterparts. Compared to ρ_α the loss function λ_α , however, is continuously differentiable, leading to favorable analytic properties in an minimization context.

3 Multivariate Geometric Risk Measures

In order to generalize univariate expectiles to the multivariate setting we first revisit in Section 3.1 the framework introduced by Chaudhuri [1996]. This allows for a suitable generalization of the loss function in (1), leading to multivariate geometric quantiles. In Section 3.2 we then apply the underlying idea of Chaudhuri [1996] to give a multivariate generalization of (2) and to introduce multivariate geometric expectiles as the main contribution of this paper.

3.1. Multivariate Geometric Value-at-Risk

Chaudhuri [1996] provides a definition of multivariate quantiles by generalizing the approach outlined in Section 2. The resulting geometric quantiles are obtained by minimizing the expected loss based on a multivariate loss function generalizing ρ_α given in (1).

For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ we denote by $\|\mathbf{x}\|_2 = \sqrt{\mathbf{x}^\top \mathbf{x}}$ and $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^\top \mathbf{y}$ the Euclidean norm and inner product respectively, and by $B = \{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\|_2 < 1\} \subset \mathbb{R}^d$ the open unit ball in \mathbb{R}^d .

For a fixed index $\mathbf{u} \in B$ Chaudhuri [1996] defines the loss function $\Phi_{\mathbf{u}}$ as

$$\Phi_{\mathbf{u}}: \mathbb{R}^d \rightarrow [0, \infty), \quad \mathbf{t} \mapsto \Phi_{\mathbf{u}}(\mathbf{t}) = \frac{1}{2}(\|\mathbf{t}\|_2 + \langle \mathbf{u}, \mathbf{t} \rangle). \quad (3)$$

While it is immediately clear that $\Phi_{\mathbf{u}}(\mathbf{0}) = 0$ for all $\mathbf{u} \in B$, we also have that $\Phi_{\mathbf{u}}(\mathbf{t}) \geq 0$ for all $(\mathbf{u}, \mathbf{t}) \in B \times \mathbb{R}^d$ using the Cauchy-Schwarz inequality. Convexity of $\Phi_{\mathbf{u}}$ follows directly from properties of the norm and inner product.

Based on $\Phi_{\mathbf{u}}$ the (*multivariate*) *geometric quantile*, or *geometric VaR*, at level $\alpha \in B$ for a random vector \mathbf{X} is then defined as

$$\text{VaR}_{\alpha}(\mathbf{X}) = \underset{\mathbf{c} \in \mathbb{R}^d}{\operatorname{argmin}} \mathbb{E}[\Phi_{\alpha}(\mathbf{X} - \mathbf{c})]. \quad (4)$$

As shown in Chaudhuri [1996], the right hand side of (4) is always finite and the minimization is thus well posed. Furthermore, the resulting geometric VaR is the unique minimizer of (4).

In (4) the vector $\alpha \in B$ takes the role of the confidence level. However, due to the multivariate context, VaR is now indexed by a d -dimensional vector instead of a scalar. This adds additional flexibility compared to other approaches such as Cousin and Di Bernardino [2013], Ben Tahar [2006] and Cossette et al. [2015], where only one scalar confidence level can be set for the multivariate risk \mathbf{X} . It is also important to notice that the geometric quantile $\text{VaR}_{\alpha}(\mathbf{X})$ itself is represented by a vector in \mathbb{R}^d . This makes the resulting risk measure easier to use for risk analysis than approaches such as Cousin and Di Bernardino [2014], Cousin and Di Bernardino [2013] and Mailhot et al. [2014] where the resulting multivariate quantiles are subsets in \mathbb{R}^d .

When comparing traditional confidence levels in $(0, 1)$ to the univariate case of our setting, care has to be taken to adjust the indices. Both settings are equivalent by simply re-indexing according to $f: [0, 1] \rightarrow [-1, 1]$, $x \mapsto f(x) = 2x - 1$. An index of 99% in the traditional setting is therefore comparable to an index of 98% using the convention adopted in this paper. The contour lines for a two dimensional example of $\Phi_{\mathbf{u}}(\mathbf{t})$ are shown in Figure 2 for three indices $\mathbf{u}_1 = 0.9/\sqrt{2}(1, 1)$, $\mathbf{u}_2 = 0.9/\sqrt{2}(-1, 1)$ and $\mathbf{u}_3 = 0.5/\sqrt{2}(-1, 1)$. The figure shows how the direction of the index, visualized by the arrow, changes the orientation of the contour lines (compare the first and second plots). On the other hand the magnitude of the index changes the shape of the contour lines (compare the second and third plots). For smaller values of $\|\mathbf{u}\|_2$ the contour lines are more norm like, i.e., more circular, and in the limit $\|\mathbf{u}\|_2 = 0$, i.e., if and only if $\mathbf{u} = (0, 0)$ we are indeed left with the circular contour lines of the norm.

3.2. Multivariate Geometric Expectiles

Analogously to the approach of Chaudhuri [1996] we introduce our multivariate representation of expectiles via a multivariate generalization of λ_{α} . For this purpose it is more convenient to rewrite the original definition of λ_{α} given in (2) as

$$\lambda_{\alpha}(t) = \frac{1}{2} |t| (|t| + (2\alpha - 1)t).$$

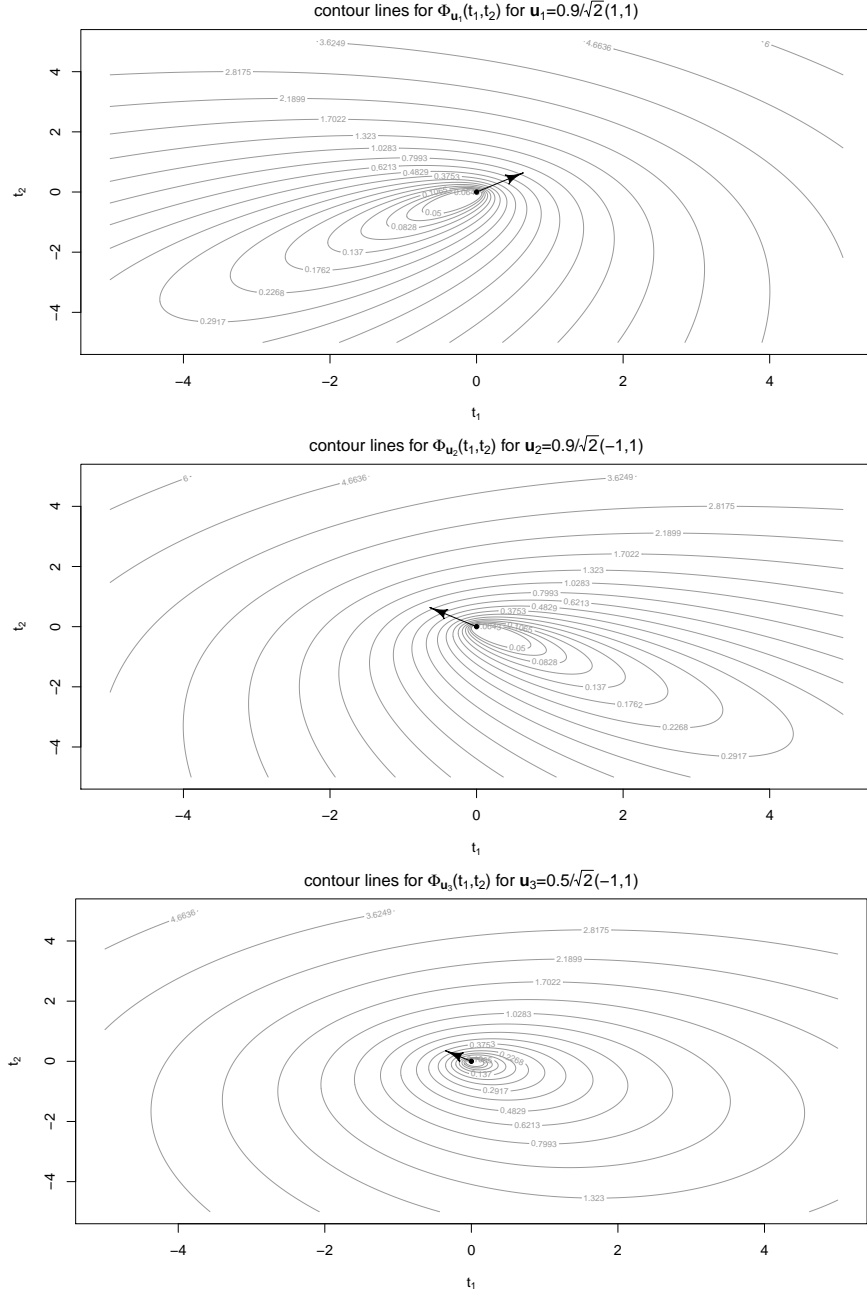


Figure 2: Contour lines for $\Phi_{\mathbf{u}_i}(t_1, t_2)$, $i \in \{1, 2, 3\}$, for indices $\mathbf{u}_1 = 0.9/\sqrt{2}(1, 1)$, $\mathbf{u}_2 = 0.9/\sqrt{2}(-1, 1)$ and $\mathbf{u}_3 = 0.5/\sqrt{2}(-1, 1)$. The global minimum is marked with a black dot at $(0, 0)$, while the arrow visualizes direction and magnitude of the index \mathbf{u}_i , $i \in \{1, 2, 3\}$.

It can easily be verified that both definitions coincide for all $t \in \mathbb{R}$. Similarly to (3) this motivates our definition of the loss function $\Lambda_{\mathbf{u}}$ as

$$\Lambda_{\mathbf{u}}: \mathbb{R}^d \rightarrow [0, \infty), \quad \mathbf{t} \mapsto \Lambda_{\mathbf{u}}(\mathbf{t}) = \frac{1}{2} \|\mathbf{t}\|_2 (\|\mathbf{t}\|_2 + \langle \mathbf{u}, \mathbf{t} \rangle), \quad (5)$$

where $\mathbf{u} \in B$ is a fixed element of the open unit ball. Given that $\Phi_{\mathbf{u}}(\mathbf{t}) \geq 0$ for all $(\mathbf{u}, \mathbf{t}) \in B \times \mathbb{R}^d$ it is clear that we also have $\Lambda_{\mathbf{u}}(\mathbf{t}) \geq 0$ for all $(\mathbf{u}, \mathbf{t}) \in B \times \mathbb{R}^d$. As for $\Phi_{\mathbf{u}}$ we have $\Lambda_{\mathbf{u}}(\mathbf{0}) = 0$ for all $\mathbf{u} \in B$.

For a given confidence level $\alpha \in B$ we now define the *geometric expectile* of a random vector \mathbf{X} as the minimizer of the expected loss based on Λ_{α} , i.e.

$$e_{\alpha}(\mathbf{X}) = \operatorname{argmin}_{\mathbf{c} \in \mathbb{R}^d} \mathbb{E}[\Lambda_{\alpha}(\mathbf{X} - \mathbf{c})]. \quad (6)$$

As in the case of the geometric VaR, the definition of geometric expectiles is based on an index $\alpha \in B$ allowing to specify a direction and magnitude of the confidence level. Furthermore, geometric expectiles are vectors in \mathbb{R}^d . This makes them easier to interpret than multivariate risk measures that are given as subsets of \mathbb{R}^d . For example, for $\alpha = \mathbf{0}$ it is easy to see that $e_{\mathbf{0}}(\mathbf{X}) = (\mathbb{E}[X_1], \dots, \mathbb{E}[X_d])$. The mean vector is therefore, analogously to the univariate case, a special case of the geometric expectiles defined in (6). In Section 4 we discuss the existence of a minimizer e_{α} and its uniqueness together with further properties of e_{α} .

Figure 3 displays the contour lines for a two dimensional example of $\Lambda_{\mathbf{u}}(\mathbf{t})$ for the three indices \mathbf{u}_1 , \mathbf{u}_2 and \mathbf{u}_3 as before. As in the case of geometric quantiles, the direction of the index, visualized by the arrow, changes the orientation of the contour lines. Also, the magnitude of the index $\|\mathbf{u}\|_2$ influences the shape of the contour lines, where smaller values of $\|\mathbf{u}\|_2$ lead to more norm like contours as already discussed in the case of quantiles. Comparing the contour lines of quantiles in Figure 2 to those of expectiles in Figure 3 also shows that the objective functions are differently shaped, more prominently so for higher values of $\|\mathbf{u}\|_2$.

The examples in Section 5 provide numerical illustrations of the resulting expectiles for a number of bivariate distributions, see Figures 5 and 8, as well as an analytic solution to (6) in the special case of a bivariate uniform distribution.

4 Properties of Geometric Expectiles

In this section we discuss properties of geometric expectiles e_{α} defined in (6). Clearly, properties of the associated loss function $\Lambda_{\mathbf{u}}$ play a major role in this discussion which is why we discuss them first in Section 4.1. In Section 4.2 we then derive properties of e_{α} . Finally, we discuss asymptotics in Section 4.3 when e_{α} needs to be estimated from observed data or approximated when closed-form solutions to the minimization problem cannot be obtained.

4.1. Properties of $\Lambda_{\mathbf{u}}$

In the univariate setting an advantage of expectiles over quantiles is that the underlying loss function is differentiable at zero. This is also true for geometric quantiles and expectiles when $d \geq 2$. The following theorem shows that the geometric expectile loss function continues to

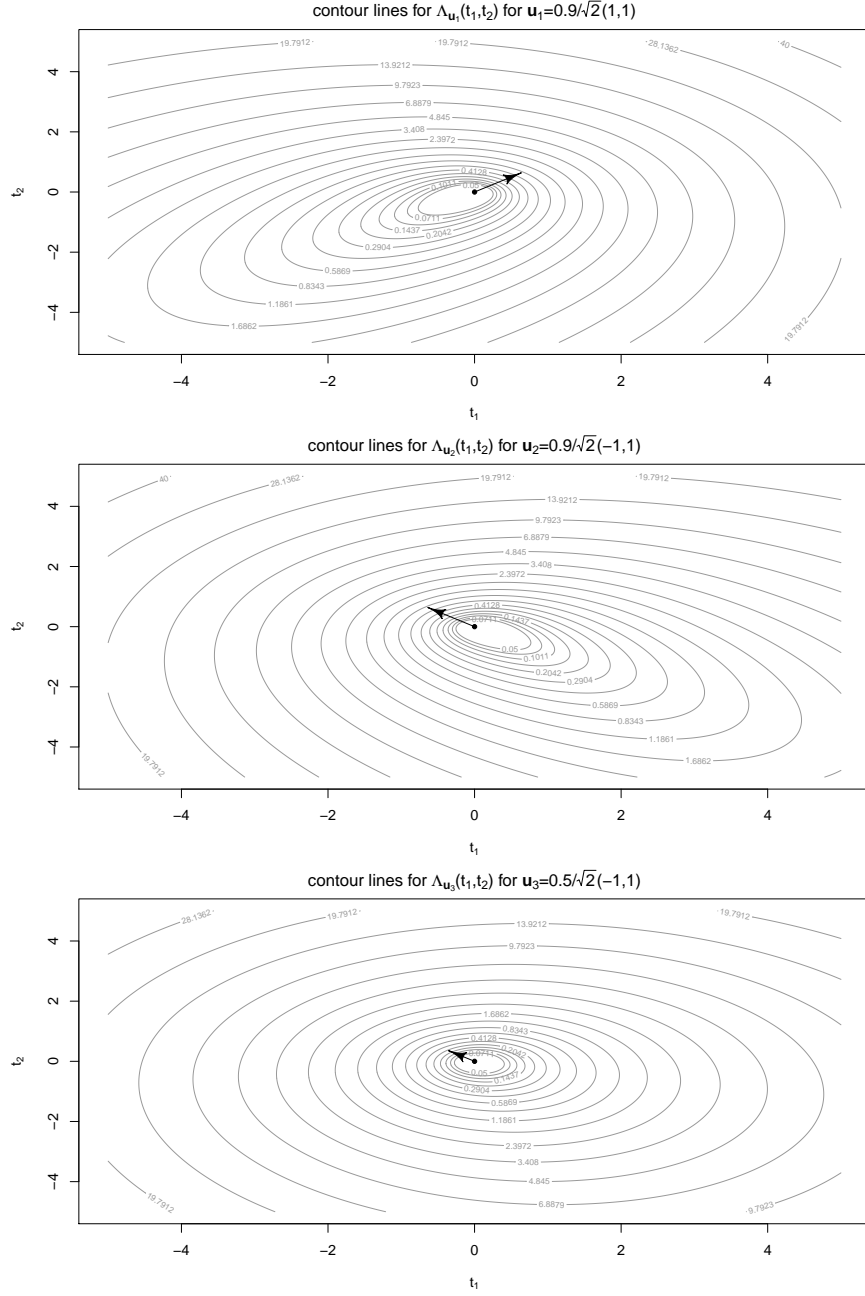


Figure 3: Contour lines for $\Lambda_{\mathbf{u}_i}(t_1, t_2)$, $i \in \{1, 2, 3\}$ for indices $\mathbf{u}_1 = 0.9/\sqrt{2}(1, 1)$, $\mathbf{u}_2 = 0.9/\sqrt{2}(-1, 1)$ and $\mathbf{u}_3 = 0.5/\sqrt{2}(-1, 1)$. The global minimum is marked with a black dot at $(0, 0)$, while the arrow visualizes direction and magnitude of the index \mathbf{u}_i , $i \in \{1, 2, 3\}$.

be differentiable for $d \geq 2$, while it is straightforward to see that this is not the case for the geometric quantile loss function $\Phi_{\mathbf{u}}$ defined in (3).

Theorem 4.1 (Differentiability of $\Lambda_{\mathbf{u}}$). *For $\Lambda_{\mathbf{u}}$ defined in (5) the gradient $\nabla \Lambda_{\mathbf{u}}(\mathbf{t})$ exists for all $(\mathbf{u}, \mathbf{t}) \in B \times \mathbb{R}^d$ with $\Lambda_{\mathbf{u}}(\mathbf{0}) = \mathbf{0}$.*

Proof. For $\mathbf{t} \neq \mathbf{0}$ it is clear that the partial derivatives with respect to each variable once exist and are finite. To show the claim for $\mathbf{t} = \mathbf{0}$ we first consider the k -th element of the gradient given by

$$\frac{\partial}{\partial t_k} \Lambda_{\mathbf{u}}(\mathbf{t}) = t_k + \frac{t_k}{2 \|\mathbf{t}\|_2} \langle \mathbf{u}, \mathbf{t} \rangle + \frac{1}{2} \|\mathbf{t}\|_2 u_k.$$

Now consider a sequence $(\mathbf{t}_n)_{n=1}^{\infty}$ such that $\lim_{n \rightarrow \infty} \mathbf{t}_n = \mathbf{0}$, and we can represent each element \mathbf{t}_n via d -dimensional polar coordinates, i.e., we consider a radius r_n and angles $\phi_{n,1}, \dots, \phi_{n,d-1}$ such that \mathbf{t}_n can be represented by

$$\mathbf{t}_n = r_n \begin{pmatrix} \cos(\phi_{n,1}) \\ \sin(\phi_{n,1}) \cos(\phi_{n,2}) \\ \sin(\phi_{n,1}) \sin(\phi_{n,2}) \cos(\phi_{n,3}) \\ \vdots \\ \sin(\phi_{n,1}) \cdots \sin(\phi_{n,d-2}) \cos(\phi_{n,d-1}) \\ \sin(\phi_{n,1}) \cdots \sin(\phi_{n,d-2}) \sin(\phi_{n,d-1}) \end{pmatrix} = r_n \boldsymbol{\xi}(\phi_{n,1}, \dots, \phi_{n,d-1})$$

where $r_n \rightarrow 0$ as $n \rightarrow \infty$. Writing ξ_1, \dots, ξ_d for the components of $\boldsymbol{\xi} = \boldsymbol{\xi}(\phi_{n,1}, \dots, \phi_{n,d-1})$ and noting that $\|\boldsymbol{\xi}\|_2 = 1$ we observe that

$$\frac{\partial}{\partial t_k} \Lambda_{\mathbf{u}}(\mathbf{t}) = r_n \xi_k + \frac{r_n \xi_k}{2 r_n} r_n \langle \mathbf{u}, \boldsymbol{\xi} \rangle + \frac{1}{2} r_n \|\boldsymbol{\xi}\|_2 u_k,$$

which converges to zero for $n \rightarrow \infty$ for any sequence $(\mathbf{t}_n)_{n=1}^{\infty}$ converging to zero. \square

From the definition of $\Phi_{\mathbf{u}}$ in (3) it is straightforward to see that $\Phi_{\mathbf{u}}$ is a convex function. While this is also true for the loss function $\Lambda_{\mathbf{u}}$ tied to geometric expectiles, it is not immediately clear from (5). To simplify the discussion we first recall a well-known result from convex analysis, see for example Rudin [1976].

Lemma 4.1 (Midpoint convexity). *Denote by $f: \mathbb{R}^d \rightarrow \mathbb{R}$ a continuous function. Then f is convex if and only if it is midpoint convex, i.e.*

$$0.5f(\mathbf{t}_1) + 0.5f(\mathbf{t}_2) - f(0.5(\mathbf{t}_1 + \mathbf{t}_2)) \geq 0$$

for all $\mathbf{t}_1, \mathbf{t}_2 \in \mathbb{R}^d$.

To further prepare the result we first present a theorem generalizing the familiar parallelogram identity. While this is an essential component of our convexity proof in Theorem 4.3, the result is interesting in its own right.

Theorem 4.2 (Parallelogram Inequality). Denote by $\overline{B} = \{\mathbf{u} \in \mathbb{R}^d : \|\mathbf{u}\|_2 \leq 1\}$ the closed unit ball in \mathbb{R}^d . For any fixed vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ it holds that

$$-\|\mathbf{x} - \mathbf{y}\|_2^2 \leq 2\|\mathbf{x}\|_2 \langle \mathbf{u}, \mathbf{x} \rangle + 2\|\mathbf{y}\|_2 \langle \mathbf{u}, \mathbf{y} \rangle - \|\mathbf{x} + \mathbf{y}\|_2 \langle \mathbf{u}, \mathbf{x} + \mathbf{y} \rangle \leq \|\mathbf{x} - \mathbf{y}\|_2^2 \quad (7)$$

for all $\mathbf{u} \in \overline{B}$. For \mathbf{u} such that $\|\mathbf{u}\|_2 < 1$ equality holds in (7) if and only if $\mathbf{x} = \mathbf{y}$.

Proof. We start by considering the bounded components of (7) as a function of \mathbf{u} which can be rewritten as

$$f_{\mathbf{x}, \mathbf{y}}(\mathbf{u}) = \langle \mathbf{u}, (2\|\mathbf{x}\|_2 - \|\mathbf{x} + \mathbf{y}\|_2)\mathbf{x} + (2\|\mathbf{y}\|_2 - \|\mathbf{x} + \mathbf{y}\|_2)\mathbf{y} \rangle.$$

First, we consider two special cases. For $\mathbf{x} = \mathbf{y}$ all terms on all sides in (7) vanish and equality holds for all $\mathbf{u} \in \overline{B}$. For $\mathbf{x} = \mathbf{0} \neq \mathbf{y}$ we have $f_{\mathbf{0}, \mathbf{y}}(\mathbf{u}) = \|\mathbf{y}\|_2 \langle \mathbf{u}, \mathbf{y} \rangle$. Therefore

$$-\|\mathbf{y}\|_2^2 \leq f_{\mathbf{0}, \mathbf{y}}(\mathbf{u}) \leq \|\mathbf{y}\|_2^2$$

holds as a consequence of the Cauchy-Schwarz inequality. Equality can only hold if $\|\mathbf{u}\|_2 = 1$. Then, for the general case, we consider $\mathbf{y} \neq \mathbf{x} \neq \mathbf{0}$ and define

$$L(\mathbf{x}, \mathbf{y}) = \|(2\|\mathbf{x}\|_2 - \|\mathbf{x} + \mathbf{y}\|_2)\mathbf{x} + (2\|\mathbf{y}\|_2 - \|\mathbf{x} + \mathbf{y}\|_2)\mathbf{y}\|_2.$$

With the Cauchy-Schwarz inequality and $\|\mathbf{u}\|_2 \leq 1$ we have that

$$-L(\mathbf{x}, \mathbf{y}) \leq f_{\mathbf{x}, \mathbf{y}}(\mathbf{u}) \leq L(\mathbf{x}, \mathbf{y}),$$

where the equality only holds if $\|\mathbf{u}\|_2 = 1$. Our claim is now equivalent to $L(\mathbf{x}, \mathbf{y}) \leq \|\mathbf{x} - \mathbf{y}\|_2^2$, or equivalently to

$$\|\mathbf{x} - \mathbf{y}\|_2^4 - L(\mathbf{x}, \mathbf{y})^2 \geq 0. \quad (8)$$

Due to the scale invariance of both terms

$$\begin{aligned} L(\sigma\mathbf{x}, \sigma\mathbf{y}) &= \sigma^2 L(\mathbf{x}, \mathbf{y}), \\ \|\sigma\mathbf{x} - \sigma\mathbf{y}\|_2^2 &= \sigma^2 \|\mathbf{x} - \mathbf{y}\|_2^2, \end{aligned}$$

for any $\sigma > 0$, we consider without loss of generality \mathbf{x} and \mathbf{y} such that $\|\mathbf{x} + \mathbf{y}\|_2 = 1$. Any other cases can be handled by rescaling with $\sigma = 1/\|\mathbf{x} + \mathbf{y}\|_2$. We continue by considering polar coordinates of $(\|\mathbf{x}\|_2, \|\mathbf{y}\|_2) \in \mathbb{R}^2$ leading to $\|\mathbf{x}\|_2 = r \cos(\theta)$ and $\|\mathbf{y}\|_2 = r \sin(\theta)$, where $r > 0$ and $0 \leq \theta \leq \pi/2$ due to the strict component wise positivity of $(\|\mathbf{x}\|_2, \|\mathbf{y}\|_2)$. This yields

$$\|\mathbf{x} + \mathbf{y}\|_2^2 = \|\mathbf{x}\|_2^2 + \|\mathbf{y}\|_2^2 + 2\langle \mathbf{x}, \mathbf{y} \rangle = r^2 + 2\langle \mathbf{x}, \mathbf{y} \rangle = 1,$$

or alternatively $(2\langle \mathbf{x}, \mathbf{y} \rangle = 1 - r^2)$. For $\|\mathbf{x} - \mathbf{y}\|_2^2$ we have

$$\|\mathbf{x} - \mathbf{y}\|_2^2 = \|\mathbf{x}\|_2^2 + \|\mathbf{y}\|_2^2 - 2\langle \mathbf{x}, \mathbf{y} \rangle = r^2 - (1 - r^2) = 2r^2 - 1.$$

For the first term in (8) we therefore have $\|\mathbf{x} - \mathbf{y}\|_2^4 = (2r^2 - 1)^2$. Concerning $L(\mathbf{x}, \mathbf{y})^2$ we have $L(\mathbf{x}, \mathbf{y})^2 = \alpha^2 \|\mathbf{x}\|_2^2 + \beta^2 \|\mathbf{y}\|_2^2 + \alpha\beta 2 \langle \mathbf{x}, \mathbf{y} \rangle$ with $\alpha = (2\|\mathbf{x}\|_2 - \|\mathbf{x} + \mathbf{y}\|_2)$ and $\beta = (2\|\mathbf{y}\|_2 - \|\mathbf{x} + \mathbf{y}\|_2)$. In terms of r and θ we then have

$$\begin{aligned}\alpha^2 &= (2r \cos(\theta) - 1)^2 \\ \beta^2 &= (2r \sin(\theta) - 1)^2 \\ \alpha\beta &= (2r \cos(\theta) - 1)(2r \sin(\theta) - 1) \\ L(\mathbf{x}, \mathbf{y})^2 &= \alpha^2 r^2 \cos(\theta)^2 + \beta^2 r^2 \sin(\theta)^2 + \alpha\beta(1 - r^2).\end{aligned}$$

Reformulating (8) in terms of r and θ then yields to

$$\|\mathbf{x} - \mathbf{y}\|_2^4 - L(\mathbf{x}, \mathbf{y})^2 = 2r(r \cos(\theta) + r \sin(\theta) - 1)^2(2r \sin(\theta) \cos(\theta) + \sin(\theta) + \cos(\theta))$$

which is non-negative given the restrictions on θ . \square

Given Lemma 4.1 and Theorem 4.2 we can now establish the strict convexity of $\Lambda_{\mathbf{u}}$.

Theorem 4.3 (Strict convexity of $\Lambda_{\mathbf{u}}$). *For every fixed $\mathbf{u} \in B$ the function $\Lambda_{\mathbf{u}}$ defined in (3) is strictly convex on \mathbb{R}^d .*

Proof. Due to continuity of $\Lambda_{\mathbf{u}}$ and Lemma 4.1 we focus on midpoint convexity. To this end, define $D(\mathbf{x}, \mathbf{y}) = 0.5\Lambda_{\mathbf{u}}(\mathbf{x}) + 0.5\Lambda_{\mathbf{u}}(\mathbf{y}) - \Lambda_{\mathbf{u}}(0.5(\mathbf{x} + \mathbf{y}))$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$, where we have that $\Lambda_{\mathbf{u}}(0.5(\mathbf{x} + \mathbf{y})) = 0.25 \|\mathbf{x} + \mathbf{y}\|_2^2 + 0.25 \|\mathbf{x} + \mathbf{y}\|_2 \langle \mathbf{u}, \mathbf{x} + \mathbf{y} \rangle = 0.25\Lambda_{\mathbf{u}}(\mathbf{x} + \mathbf{y})$. The function $\Lambda_{\mathbf{u}}$ is then convex if and only if $h: \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$, $(\mathbf{x}, \mathbf{y}) \mapsto h(\mathbf{x}, \mathbf{y}) = 4D(\mathbf{x}, \mathbf{y}) = 2\Lambda_{\mathbf{u}}(\mathbf{x}) + 2\Lambda_{\mathbf{u}}(\mathbf{y}) - \Lambda_{\mathbf{u}}(\mathbf{x} + \mathbf{y})$ is non-negative. For h we have that

$$\begin{aligned}h(\mathbf{x}, \mathbf{y}) &= 2\|\mathbf{x}\|_2^2 + 2\|\mathbf{x}\|_2 \langle \mathbf{u}, \mathbf{x} \rangle + 2\|\mathbf{y}\|_2^2 + 2\|\mathbf{y}\|_2 \langle \mathbf{u}, \mathbf{y} \rangle - \|\mathbf{x} + \mathbf{y}\|_2^2 - \|\mathbf{x} + \mathbf{y}\|_2 \langle \mathbf{u}, \mathbf{x} + \mathbf{y} \rangle \\ &= \|\mathbf{x} - \mathbf{y}\|_2^2 + 2\|\mathbf{x}\|_2 \langle \mathbf{u}, \mathbf{x} \rangle + 2\|\mathbf{y}\|_2 \langle \mathbf{u}, \mathbf{y} \rangle - \|\mathbf{x} + \mathbf{y}\|_2 \langle \mathbf{u}, \mathbf{x} + \mathbf{y} \rangle,\end{aligned}$$

where we used the parallelogram identity $2\|\mathbf{x}\|_2^2 + 2\|\mathbf{y}\|_2^2 = \|\mathbf{x} + \mathbf{y}\|_2^2 + \|\mathbf{x} - \mathbf{y}\|_2^2$ to get the second equality. The condition $h(\mathbf{x}, \mathbf{y}) \geq 0$ is equivalent to

$$-\|\mathbf{x} - \mathbf{y}\|_2^2 \leq 2\|\mathbf{x}\|_2 \langle \mathbf{u}, \mathbf{x} \rangle + 2\|\mathbf{y}\|_2 \langle \mathbf{u}, \mathbf{y} \rangle - \|\mathbf{x} + \mathbf{y}\|_2 \langle \mathbf{u}, \mathbf{x} + \mathbf{y} \rangle,$$

which holds by Theorem 4.2. The fact that $\Lambda_{\mathbf{u}}$ is strictly convex follows, as the index \mathbf{u} is assumed to lie in the open ball, i.e., $\|\mathbf{u}\|_2 = 1$ is not permitted. \square

We have so far established that $\Lambda_{\mathbf{u}}$ is differentiable with a stationary point at $\mathbf{t} = \mathbf{0}$. Furthermore, the strict convexity of $\Lambda_{\mathbf{u}}$ guarantees that there exists at most one global minimum for $\Lambda_{\mathbf{u}}$. To finally ensure the existence of such a minimizer we establish coercivity of $\Lambda_{\mathbf{u}}$.

Definition 4.1 (Coercive function on \mathbb{R}^d). *A real valued function $f: \mathbb{R}^d \rightarrow \mathbb{R}$ is said to be coercive if $\lim_{n \rightarrow \infty} f(\mathbf{x}_n) = \infty$ for all sequences $(\mathbf{x}_n)_{n=1}^{\infty}$ such that $\lim_{n \rightarrow \infty} \|\mathbf{x}_n\|_2 = \infty$.*

Coercivity plays an important role in optimization theory as it ensures the existence of at least one minimizer for a large class of real valued functions. This fact is formalized in the following theorem.

Theorem 4.4. Denote by $f: \mathbb{R}^d \rightarrow \mathbb{R}$ a coercive and convex function. Then there exists an element $\mathbf{x}_0 \in \mathbb{R}^d$ such that $f(\mathbf{x}_0) = \inf_{\mathbf{x} \in \mathbb{R}^d} f(\mathbf{x})$.

Proof. The proof follows from the more general Theorem 2.11 and Remark 2.13 in Barbu and Precupanu [2012] applicable to lower-semicontinuous functions on reflexive Banach spaces. The necessary continuity of f is guaranteed by Proposition 2.3 in Tuy [2016] stating that a proper convex function on \mathbb{R}^d is continuous on every interior point of its effective domain. \square

To finally tie all parts together we establish the coercivity of $\Lambda_{\mathbf{u}}$ in the following theorem. Given the strict convexity of $\Lambda_{\mathbf{u}}$ established in Theorem 4.3, an application of Theorem 4.4 ensures the existence of a unique and global minimizer. From our previous observations, especially Theorem 4.1, we know that this minimum is located at $\mathbf{0}$ for every given $\mathbf{u} \in B$.

Theorem 4.5 (Coercivity of $\Lambda_{\mathbf{u}}$). *The function $\Lambda_{\mathbf{u}}$ is coercive on \mathbb{R}^d .*

Proof. Given that $\|\mathbf{u}\|_2 = s < 1$ the Cauchy-Schwarz inequality implies $\langle \mathbf{u}, \mathbf{x} \rangle \geq -s \|\mathbf{x}\|_2$. Therefore $\Lambda_{\mathbf{u}}(\mathbf{x}) \geq 0.5 \|\mathbf{x}\|_2^2 (1 - s)$ which proves the claim. \square

4.2. Properties of e_{α}

With the properties of $\Lambda_{\mathbf{u}}$ in place we can now tackle those of e_{α} . In (6), it is necessary to ensure that the objective function, i.e., the expected loss, is finite. Similarly to the univariate case, we recover that a finite second moment condition for the marginal distributions is sufficient. We thus introduce the condition

(C) For a d -dimensional random vector \mathbf{X} assume $\mathbb{E}[X_j^2] < \infty$ for all $j \in \{1, \dots, d\}$

for ease of reference. This leads to the following result.

Theorem 4.6. *If (C) holds for a d -dimensional random vector $\mathbf{X} = (X_1, \dots, X_d)$ then $0 \leq \mathbb{E}[\Lambda_{\mathbf{u}}(\mathbf{X} - \mathbf{c})] < \infty$ for every $\mathbf{c} \in \mathbb{R}^d$ and $\mathbf{u} \in B$.*

Proof. We use Jensen's inequality and $\|\mathbf{u}\|_2 < 1$ to obtain that

$$|\mathbb{E}[\|\mathbf{X}\|_2 \langle \mathbf{u}, \mathbf{X} \rangle]| \leq \mathbb{E}[\|\mathbf{X}\|_2 |\langle \mathbf{u}, \mathbf{X} \rangle|] \leq \mathbb{E}[\|\mathbf{X}\|_2^2].$$

For a given $\mathbf{c} \in \mathbb{R}^d$ and $\mathbf{u} \in B$, this leads to

$$\begin{aligned} \mathbb{E}[\Lambda_{\mathbf{u}}(\mathbf{X} - \mathbf{c})] &= |\mathbb{E}[\Lambda_{\mathbf{u}}(\mathbf{X} - \mathbf{c})]| \leq \mathbb{E}[\|\mathbf{X} - \mathbf{c}\|_2^2] + \mathbb{E}[\|\mathbf{X} - \mathbf{c}\|_2 |\langle \mathbf{u}, \mathbf{X} - \mathbf{c} \rangle|] \\ &\leq 2\mathbb{E}[\|\mathbf{X} - \mathbf{c}\|_2^2] \\ &= 2 \sum_{j=1}^d \mathbb{E}[(X_j - c_j)^2] < \infty, \end{aligned}$$

as $\Lambda_{\mathbf{u}}(\mathbf{t}) \geq 0$ for all $(\mathbf{u}, \mathbf{t}) \in B \times \mathbb{R}^d$ and $\mathbb{E}[X_j^2] < \infty$. \square

Now that the finiteness of the expected loss is addressed, we turn to the existence and uniqueness of e_{α} . To do so we adapt the proof of Theorem 6.8 in Lehmann [1983] to our more general setting. To this end let

$$\phi(\mathbf{c}) = \mathbb{E}[\Lambda_{\alpha}(\mathbf{X} - \mathbf{c})] \tag{9}$$

denote the objective function used in (6) and recall the convergence in probability to infinity.

Definition 4.2 (Convergence in probability to ∞). *A sequence of positive random variables $(Y_n)_{n=1}^\infty$ converges in probability to ∞ , if, for every $K > 0$, $\lim_{n \rightarrow \infty} \mathbb{P}[Y_n > K] = 1$.*

In preparation of showing coercivity of ϕ , we first discuss the probabilistic behaviour of $\Lambda_\alpha(\mathbf{X} - \mathbf{c})$ when $\|\mathbf{c}\|_2$ tends towards ∞ .

Lemma 4.2. *For a sequence of vectors $(\mathbf{c}_n)_{n=1}^\infty$ such that $\|\mathbf{c}_n\|_2 \rightarrow \infty$ and for a fixed random vector \mathbf{X} , the sequence $(\Lambda_\alpha(\mathbf{X} - \mathbf{c}_n))_{n=1}^\infty$ converges in probability to ∞ .*

Proof. From the proof of Theorem 4.5 we have that $\Lambda_{\mathbf{u}}(\mathbf{t}) \geq 0.5(1-s) \|\mathbf{t}\|_2^2$, where $s = \|\mathbf{u}\|_2$. Therefore, almost surely,

$$\Lambda_{\mathbf{u}}(\mathbf{X} - \mathbf{c}) \geq 0.5(1-s) \|\mathbf{X} - \mathbf{c}\|_2^2. \quad (10)$$

With the reverse triangle inequality, we also have that

$$0.5(1-s) \left| \|\mathbf{X}\|_2 - \|\mathbf{c}\|_2 \right|^2 \leq 0.5(1-s) \|\mathbf{X} - \mathbf{c}\|_2^2, \quad (11)$$

almost surely. For an arbitrary fixed $K > 0$, we define the sets

$$\begin{aligned} A_n(K) &= \{\omega \in \Omega : 0.5(1-s) \left| \|\mathbf{X}\|_2 - \|\mathbf{c}_n\|_2 \right|^2 > K\}, \\ B_n(K) &= \{\omega \in \Omega : 0.5(1-s) \|\mathbf{X} - \mathbf{c}_n\|_2^2 > K\}, \\ C_n(K) &= \{\omega \in \Omega : \Lambda_{\mathbf{u}}(\mathbf{X} - \mathbf{c}_n) > K\}, \end{aligned}$$

leading to $A_n(K) \subset B_n(K) \subset C_n(K)$ on the basis of inequalities (10) and (11). For $A_n(K)$ and every $K > 0$, we have that

$$\begin{aligned} \mathbb{P}[A_n(K)] &= 1 - \mathbb{P}[0.5(1-s) \left| \|\mathbf{X}\|_2 - \|\mathbf{c}_n\|_2 \right|^2 \leq K], \\ &= 1 - \mathbb{P} \left[\left| \|\mathbf{X}\|_2 - \|\mathbf{c}_n\|_2 \right| \leq \sqrt{\frac{2K}{1-s}} \right], \\ &= 1 - \left(F_{\|\mathbf{X}\|_2} \left(\|\mathbf{c}_n\|_2 + \sqrt{\frac{2K}{1-s}} \right) - F_{\|\mathbf{X}\|_2} \left(\|\mathbf{c}_n\|_2 - \sqrt{\frac{2K}{1-s}} \right) \right) \rightarrow 1, \end{aligned}$$

for $n \rightarrow \infty$. Since $\mathbb{P}[A_n(K)] \leq \mathbb{P}[B_n(K)] \leq \mathbb{P}[C_n(K)]$, $\lim_{n \rightarrow \infty} \mathbb{P}[C_n(K)] = 1$ for every $K > 0$. \square

In a second step, we show that the strict convexity of $\Lambda_{\mathbf{u}}$ established in Theorem 4.3 carries over to ϕ .

Theorem 4.7 (Strict convexity and continuity of ϕ). *If (C) holds for a random vector \mathbf{X} then $\phi: \mathbb{R}^d \rightarrow [0, \infty)$, $\mathbf{c} \mapsto \phi(\mathbf{c}) = \mathbb{E}[\Lambda_\alpha(\mathbf{X} - \mathbf{c})]$ is strictly convex and continuous on \mathbb{R}^d for every fixed $\alpha \in B$.*

Proof. Given that the marginal second moments of \mathbf{X} are finite, Theorem 4.6 guarantees that ϕ is well-defined for every $\mathbf{c} \in \mathbb{R}^d$. With $\lambda \in [0, 1]$ and $\mathbf{c}_1, \mathbf{c}_2 \in \mathbb{R}^d$, we have that

$$\begin{aligned} \phi(\lambda \mathbf{c}_1 + (1-\lambda)\mathbf{c}_2) &= \mathbb{E}[\Lambda_\alpha(\mathbf{X} - (\lambda \mathbf{c}_1 + (1-\lambda)\mathbf{c}_2))], \\ &= \mathbb{E}[\Lambda_\alpha(\lambda \mathbf{X} + (1-\lambda)\mathbf{X} - \lambda \mathbf{c}_1 - (1-\lambda)\mathbf{c}_2)], \\ &= \mathbb{E}[\Lambda_\alpha(\lambda(\mathbf{X} - \mathbf{c}_1) + (1-\lambda)(\mathbf{X} - \mathbf{c}_2))], \\ &< \mathbb{E}[\lambda \Lambda_\alpha(\mathbf{X} - \mathbf{c}_1) + (1-\lambda) \Lambda_\alpha(\mathbf{X} - \mathbf{c}_2)], \\ &= \lambda \phi(\mathbf{c}_1) + (1-\lambda) \phi(\mathbf{c}_2). \end{aligned}$$

Continuity follows from the fact that every proper convex function on \mathbb{R}^d is continuous on every interior point of its effective domain, see, for example, Proposition 2.3 of Tuy [2016]. Having established that ϕ is strictly convex on all of \mathbb{R}^d the claim follows. \square

Combining Lemma 4.2 and Theorem 4.7 now allows us to ensure the existence and uniqueness of geometric expectiles.

Theorem 4.8 (Existence and uniqueness of e_α). *If (C) holds for a d -dimensional random vector \mathbf{X} then there exists a unique solution $e_\alpha(\mathbf{X}) = \operatorname{argmin}_{\mathbf{c} \in \mathbb{R}^d} \phi(\mathbf{c})$ for every fixed $\alpha \in B$.*

Proof. First, we show that ϕ is coercive and fix a sequence $(\mathbf{c}_n)_{n=1}^\infty$ such that $\|\mathbf{c}_n\|_2 \rightarrow \infty$. To show that $\phi(\mathbf{c}_n)$ diverges, we fix an arbitrary $K > 0$ and define

$$C_n(K) = \{\omega \in \Omega : \Lambda(\alpha, \mathbf{X} - \mathbf{c}_n) > K\}.$$

Given that Λ_α is positive, it follows that

$$\begin{aligned} \phi(\mathbf{c}_n) &= \mathbb{E}[\Lambda_\alpha(\mathbf{X} - \mathbf{c}_n)] = \int_{C_n(K)} \Lambda_\alpha(\mathbf{X} - \mathbf{c}_n) d\mathbb{P} + \int_{\Omega \setminus C_n(K)} \Lambda_\alpha(\mathbf{X} - \mathbf{c}_n) d\mathbb{P} \\ &\geq \int_{C_n(K)} \Lambda_\alpha(\mathbf{X} - \mathbf{c}_n) d\mathbb{P} > K \mathbb{P}[C_n(K)]. \end{aligned}$$

Lemma 4.2 yields $K \mathbb{P}[C_n(K)] \rightarrow K$ as $n \rightarrow \infty$, which shows that ϕ diverges, i.e., that ϕ is coercive. Note that from Theorem 4.7, ϕ is also continuous and strictly convex. Then apply Theorem 4.4. \square

Having established the basic properties of geometric expectiles, we now discuss their behaviour under data transformations. As in Chaudhuri [1996] for $\operatorname{VaR}_\alpha(\mathbf{X})$, it is straightforward to show how geometric expectiles behave for translation, rotation and rescaling of the underlying random vector \mathbf{X} . Adding a deterministic amount to an uncertain position simply shifts the resulting risk measure, in line with translation invariance of coherent risk measures.

Proposition 4.1 (Translation invariance). *If (C) holds for a d -dimensional random vector \mathbf{X} then $e_\alpha(\mathbf{X} + \mathbf{a}) = e_\alpha(\mathbf{X}) + \mathbf{a}$ for all $\mathbf{a} \in \mathbb{R}^d$.*

Proof. By definition, we have $e_\alpha(\mathbf{X}) = \operatorname{argmin}_{\mathbf{c} \in \mathbb{R}^d} \mathbb{E}[\Lambda_\alpha(\mathbf{X} - \mathbf{c})]$. Therefore, $\mathbb{E}[\Lambda_\alpha(\mathbf{X} - (\mathbf{c} - \mathbf{a}))]$ will be minimized by $e_\alpha(\mathbf{X}) + \mathbf{a}$. \square

Reasonable behaviour under scaling transformations ensures that a change in the underlying measurement units (for example, going from cents to dollars) is appropriately reflected in the behaviour of the risk measure. For geometric expectiles this is the case as shown next. This is in resemblance with positive homogeneity of coherent risk measures.

Proposition 4.2 (Positive homogeneity). *If (C) holds for a d -dimensional random vector \mathbf{X} then $e_\alpha(\sigma \mathbf{X}) = \sigma e_\alpha(\mathbf{X})$ for every positive scalar $\sigma > 0$.*

Proof.

$$\begin{aligned}
\Lambda_{\alpha}(\sigma \mathbf{X} - \mathbf{c}) &= \|\sigma \mathbf{X} - \mathbf{c}\|_2 (\|\sigma \mathbf{X} - \mathbf{c}\|_2 + \langle \alpha, \sigma \mathbf{X} - \mathbf{c} \rangle) \\
&= \sigma \|\mathbf{X} - \sigma^{-1} \mathbf{c}\|_2 (\sigma \|\mathbf{X} - \sigma^{-1} \mathbf{c}\|_2 + \sigma \langle \alpha, \mathbf{X} - \sigma^{-1} \mathbf{c} \rangle) \\
&= \sigma^2 \Lambda_{\alpha}(\mathbf{X} - \sigma^{-1} \mathbf{c}).
\end{aligned}$$

Given that $e_{\alpha}(\mathbf{X})$ minimizes $\sigma^2 \mathbb{E}[\Lambda_{\alpha}(\mathbf{X} - \mathbf{c})]$, as the positive factor σ^2 only changes the value of the objective function but not the location of the optimum, we have that $\sigma e_{\alpha}(\mathbf{X})$ minimizes $\mathbb{E}[\Lambda_{\alpha}(\sigma \mathbf{X} - \mathbf{c})]$, i.e., $e_{\alpha}(\sigma \mathbf{X}) = \sigma e_{\alpha}(\mathbf{X})$. \square

It is reasonable to expect that a permutation of the components of \mathbf{X} should likewise result in a permutation of the entries of the resulting risk measure. Geometric expectiles are not only well behaved under permutations, but under general orthogonal rotations.

Proposition 4.3 (Rotation with orthogonal matrix). *If (C) holds for a d -dimensional random vector \mathbf{X} then $e_{A\alpha}(A\mathbf{X}) = Ae_{\alpha}(\mathbf{X})$ for every orthogonal matrix $A \in \mathbb{R}^{d \times d}$.*

Proof. By orthogonality of A , for every $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$, we have that $\langle A\mathbf{x}, A\mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$ and $\|A\mathbf{x}\|_2 = \|\mathbf{x}\|_2$. Therefore, we get

$$\begin{aligned}
\Lambda_{A\alpha}(A\mathbf{X} - \mathbf{c}) &= \|A\mathbf{X} - \mathbf{c}\|_2 (\|A\mathbf{X} - \mathbf{c}\|_2 + \langle A\alpha, A\mathbf{X} - \mathbf{c} \rangle) \\
&= \|\mathbf{X} - A^{\top} \mathbf{c}\|_2 (\|\mathbf{X} - A^{\top} \mathbf{c}\|_2 + \langle \alpha, \mathbf{X} - A^{\top} \mathbf{c} \rangle) \\
&= \Lambda_{\alpha}(\mathbf{X} - A^{\top} \mathbf{c}).
\end{aligned}$$

Given that $e_{\alpha}(\mathbf{X})$ minimizes $\mathbb{E}[\Lambda_{\alpha}(\mathbf{X} - \mathbf{c})]$, the minimizer of $\mathbb{E}[\Lambda_{\alpha}(\mathbf{X} - A^{\top} \mathbf{c})]$ is given by $Ae_{\alpha}(\mathbf{X})$. \square

Taken together, Propositions 4.1–4.3 guarantee that $e_{\alpha}(\mathbf{X})$ is well behaved for the most relevant data transformations. In this context it is natural to ask if there exists a suitable ordering \prec for random vectors such that $\mathbf{X} \prec \mathbf{Y}$ implies $e_{\alpha}(\mathbf{X}) \sqsubset e_{\alpha}(\mathbf{Y})$ for a possibly different ordering \sqsubset . While this point is of great interest it proved too difficult to establish a suitable result and we thus leave it as an open question for further research.

In the following Proposition 4.4 and Proposition 4.5 we generalize well known symmetry properties of univariate expectiles to the multivariate setting. We start by establishing a connection between the geometric expectiles of \mathbf{X} and $-\mathbf{X}$.

Proposition 4.4 (Vector sign symmetry). *If (C) holds for a d -dimensional random vector \mathbf{X} then $e_{\alpha}(-\mathbf{X}) = -e_{-\alpha}(\mathbf{X})$ for all $\alpha \in B$.*

Proof. For $\alpha \in B$ we have

$$\begin{aligned}
\mathbb{E}[\Lambda_{\alpha}(-\mathbf{X} + e_{-\alpha}(\mathbf{X}))] &= \mathbb{E}[\|-\mathbf{X} + e_{-\alpha}(\mathbf{X})\|_2^2] + \mathbb{E}[\|-\mathbf{X} + e_{-\alpha}(\mathbf{X})\|_2 \langle \alpha, -\mathbf{X} + e_{-\alpha}(\mathbf{X}) \rangle] \\
&= \mathbb{E}[\|\mathbf{X} - e_{-\alpha}(\mathbf{X})\|_2^2] + \mathbb{E}[\|\mathbf{X} - e_{-\alpha}(\mathbf{X})\|_2 \langle -\alpha, \mathbf{X} - e_{-\alpha}(\mathbf{X}) \rangle] \\
&= \mathbb{E}[\Lambda_{-\alpha}(\mathbf{X} - e_{-\alpha}(\mathbf{X}))].
\end{aligned}$$

The claim follows since the right hand side is minimized by definition of $e_{-\alpha}(\mathbf{X})$. \square

For radially symmetric distributions, see for example McNeil et al. [2015] Chapter 7, the resulting expectiles also obey a symmetry relation when changing the sign of the underlying index α .

Proposition 4.5 (Index sign symmetry). *If (C) holds for a d -dimensional radially symmetric random vector \mathbf{X} a with mean vector $\boldsymbol{\mu}$ then $e_{-\alpha}(\mathbf{X}) = 2\boldsymbol{\mu} - e_{\alpha}(\mathbf{X})$ for all $\alpha \in B$.*

Proof. For $\alpha \in B$ we have

$$\begin{aligned}\mathbb{E}[\Lambda_{-\alpha}(\mathbf{X} - \mathbf{c})] &= \mathbb{E}[\|\mathbf{X} - \mathbf{c}\|_2^2] + \mathbb{E}[\|\mathbf{X} - \mathbf{c}\|_2 \langle -\alpha, \mathbf{X} - \mathbf{c} \rangle] \\ &= \mathbb{E}[\|-(\mathbf{c} - \mathbf{X})\|_2^2] + \mathbb{E}[\|-(\mathbf{c} - \mathbf{X})\|_2 \langle \alpha, \mathbf{c} - \mathbf{X} \rangle] \\ &= \mathbb{E}[\Lambda_{\alpha}(\mathbf{c} - \mathbf{X})] \\ &= \mathbb{E}[\Lambda_{\alpha}(-(\mathbf{X} - \mathbf{c}))]\end{aligned}$$

for all $\mathbf{c} \in \mathbb{R}^d$. By translation invariance (Proposition 4.1) and $\mathbf{X} - \boldsymbol{\mu} \stackrel{d}{=} -(\mathbf{X} - \boldsymbol{\mu})$ we therefore have

$$\begin{aligned}\mathbb{E}[\Lambda_{-\alpha}(\mathbf{X} - (2\boldsymbol{\mu} - e_{\alpha}(\mathbf{X})))] &= \mathbb{E}[\Lambda_{\alpha}(-(\mathbf{X} - \boldsymbol{\mu}) - e_{\alpha}(\mathbf{X}) + \boldsymbol{\mu})] \\ &= \mathbb{E}[\Lambda_{\alpha}((\mathbf{X} - \boldsymbol{\mu}) - (e_{\alpha}(\mathbf{X}) - \boldsymbol{\mu}))] \\ &= \mathbb{E}[\Lambda_{\alpha}((\mathbf{X} - \boldsymbol{\mu}) - e_{\alpha}(\mathbf{X} - \boldsymbol{\mu}))],\end{aligned}$$

where the right hand side is minimized implying that $e_{-\alpha}(\mathbf{X}) = 2\boldsymbol{\mu} - e_{\alpha}(\mathbf{X})$. \square

In the univariate setting, expectiles are an attractive choice among possible risk measures due to their elicibility. As discussed in Gneiting [2011], elicibility is a property of statistical functionals when considering point forecasts. Denoting by \mathcal{F} the class of probability distributions on \mathbb{R}^d with finite second marginal moments, we denote by T a statistical functional, i.e.,

$$T: \mathcal{F} \rightarrow \mathbb{R}^d, \quad F \mapsto T(F).$$

Statistical functionals can, in general, be set valued maps, as for example in the case of quantiles. However, we will here concentrate on the case where they take values in the Euclidean space, and we thus adjust the definition of elicibility given in Gneiting [2011] to this case.

Definition 4.3 (Elicibility). *A statistical functional T is called elicitable relative to the class \mathcal{F} , if*

- 1) *there exists a scoring function $S: \mathbb{R}^d \times \mathbb{R}^d \rightarrow [0, \infty)$, $(\mathbf{x}, \mathbf{y}) \mapsto S(\mathbf{x}, \mathbf{y})$ such that there is a representation*

$$T(F) = \operatorname{argmin}_{\mathbf{c} \in \mathbb{R}^d} \mathbb{E}[S(\mathbf{c}, \mathbf{X})],$$

for every $F \in \mathcal{F}$ where $\mathbf{X} \sim F$, and

- 2) $\mathbb{E}[S(T(F), \mathbf{X})] = \mathbb{E}[S(\mathbf{c}, \mathbf{X})]$ *implies $\mathbf{c} = T(F)$.*

A functional T is therefore elicitable, if it can be represented as the unique minimizer of a Bayes rule for a suitable scoring function. For geometric expectiles, we can define, for $\alpha \in B$, an associated functional T_α as

$$T_\alpha: \mathcal{F} \rightarrow \mathbb{R}^d, \quad F \mapsto e_\alpha(\mathbf{X}) \text{ for } \mathbf{X} \sim F,$$

where Theorem 4.8 guarantees that $T_\alpha(F)$ is not set-valued. It is then clear from the definition of e_α that the scoring function

$$S_\alpha: \mathbb{R}^d \times \mathbb{R}^d, \quad (\mathbf{x}, \mathbf{y}) \mapsto \Lambda_\alpha(\mathbf{x} - \mathbf{y})$$

makes T_α elicitable relative to the class \mathcal{F} . Again here, Theorem 4.8 plays a crucial role under the assumption of a joint distribution with margins with finite second moments.

In the univariate case elicibility allows to assess and compare the forecasting performance of different competing models, see Nolde and Ziegel [2016] for a discussion. In a practical setting this allows one to select a best model based on expectile point-forecasting performance and to implement meaningful expectile-based backtesting procedures against real data. The elicibility of geometric expectiles now possibly opens the door to implement model selection and backtesting procedures for the underlying joint distribution as opposed to the the marginal distributions only. From a theoretical perspective, geometric expectiles also add to a further understanding of multivariate elicibility by providing a scoring function that is not a linear combination of univariate scoring functions; see Fissler and Ziegel [2016] for an in depth discussion.

The scoring function S_α tied to geometric expectiles is furthermore positive homogeneous of order two as shown in Proposition 4.6 below. Efron [1991] highlights the necessity of positive homogeneity, or scale invariance, in an estimation context. Scale invariance and estimation of scale is also central to the theory of robust statistics; see, for example, Huber and Ronchetti [2009]. Patton [2011] furthermore argues for homogeneity in the context of forecast rankings, as the rankings obtained from a homogenous scoring function are invariant to a re-scaling of the underlying data. See also Gneiting [2011] and Nolde and Ziegel [2016] for a discussion in the context of univariate expectiles.

By establishing the positive homogeneity of S_α we prepare likewise applications in the multivariate case.

Proposition 4.6 (Positive homogeneity of S_α of order 2). *For $c > 0$ and $(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^d \times \mathbb{R}^d$, $S_\alpha(c\mathbf{x}, c\mathbf{y}) = c^2 S_\alpha(\mathbf{x}, \mathbf{y})$.*

Proof. Using basic properties of norms and inner products we have that

$$\begin{aligned} S_\alpha(c\mathbf{x}, c\mathbf{y}) &= \Lambda_\alpha(c(\mathbf{x} - \mathbf{y})) = \frac{1}{2} \|c(\mathbf{x} - \mathbf{y})\|_2 (\|c(\mathbf{x} - \mathbf{y})\|_2 + \langle \mathbf{u}, c(\mathbf{x} - \mathbf{y}) \rangle) \\ &= c^2 \frac{1}{2} \|\mathbf{x} - \mathbf{y}\|_2 (\|\mathbf{x} - \mathbf{y}\|_2 + \langle \mathbf{u}, \mathbf{x} - \mathbf{y} \rangle) = c^2 \Lambda_\alpha(\mathbf{x} - \mathbf{y}) = c^2 S_\alpha(\mathbf{x}, \mathbf{y}). \quad \square \end{aligned}$$

4.3. Asymptotics and Estimation

In Section 4 we have established that geometric expectiles defined in (6) are a well-defined functional for random vectors with finite marginal second moments. In terms of practical

applications, this raises two questions. First, the computation of closed-form solutions of $e_{\alpha}(\mathbf{X})$ might not be possible for a given random vector \mathbf{X} and numerical approximation needs to be invoked instead. Second, in practical applications it is necessary to establish that a sample version of $e_{\alpha}(\mathbf{X})$ is a consistent estimator of $e_{\alpha}(\mathbf{X})$. While the implicit definition of $e_{\alpha}(\mathbf{X})$ might seem challenging at first, our functional falls into the well-established framework of M-estimators; see Huber and Ronchetti [2009] for an introduction.

To discuss consistency, we denote by $(\mathbf{X}_i)_{i=1}^{\infty}$ a sequence of independent and identically distributed (iid) random vectors with the same distribution as \mathbf{X} . While a generalization to ergodic and (weakly) stationary random vectors is straight forward, we focus on the iid case for ease of presentation. For a finite sample we replace the expectation in (6) by the sample average. This provides a finite sample version, or Monte Carlo estimator, of ϕ defined in (9) by

$$\phi_n(\mathbf{c}) = \frac{1}{n} \sum_{i=1}^n \Lambda_{\alpha}(\mathbf{X}_i - \mathbf{c}),$$

where we immediately get $\phi_n(\mathbf{c}) \xrightarrow{a.s.} \phi(\mathbf{c})$ (and thus $\phi_n(\mathbf{c}) \xrightarrow{p} \phi(\mathbf{c})$) from the strong law of large numbers. To also guarantee the convergence of the minimizers we invoke Proposition 7.4 of Hayashi [2000].

Corollary 4.1 (Consistency). *If (C) holds for a d -dimensional random vector \mathbf{X} then*

$$\operatorname{argmin}_{\mathbf{c} \in \mathbb{R}^d} \phi_n(\mathbf{c}) \xrightarrow{p} e_{\alpha}(\mathbf{X}).$$

Proof. We apply Proposition 7.4 of Hayashi [2000] which guarantees the consistency of M-estimators. By Theorem 4.8 $\phi(\mathbf{c})$ is uniquely minimized on \mathbb{R}^d at $e_{\alpha}(\mathbf{X})$, and $\phi(\mathbf{c})$ exists and is finite for all $\mathbf{c} \in \mathbb{R}^d$. Furthermore, Λ_{α} is convex. While the existence of a minimizer in Proposition 7.4 in Hayashi [2000] is only asymptotic, it is clear that a minimizer exists for every $n \in \mathbb{N}$ in our case. \square

Corollary 4.1 also suggests a simple approach to compute $e_{\alpha}(\mathbf{X})$ when a close form solution cannot be established. If a sampling method for \mathbf{X} is available then replacing the expectation by an empirical mean yields a valid approximation. Denoting by $(\mathbf{x}_n)_{i=1}^n$ a realization of a sequence of random vectors $(\mathbf{X}_i)_{i=1}^n$, either obtained by simulation or from real data, it also important to notice that ϕ_n is strictly convex.

Corollary 4.2 (Strict convexity of ϕ_n). *Denote by $(\mathbf{x}_i)_{i=1}^n$ a sequence of vectors in \mathbb{R}^d . The function $\phi_n: \mathbb{R}^d \rightarrow \mathbb{R}$, $\mathbf{c} \mapsto \phi_n(\mathbf{c}) = \frac{1}{n} \sum_{i=1}^n \Lambda_{\alpha}(\mathbf{x}_i - \mathbf{c})$ is strictly convex.*

Proof. Given that ϕ_n is a convex combination of strictly convex functions the proof follows from basic properties of convex functions. \square

The importance of Corollary 4.2 is that the minimization

$$\operatorname{argmin}_{\mathbf{c} \in \mathbb{R}^d} \phi_n(\mathbf{c})$$

is well behaved also in the finite sample case and there exists a unique minimizer that is consistent for the functional according to Corollary 4.1. The minimizer, i.e., the finite sample version of $e_{\alpha}(\mathbf{X})$, can then be obtained by numerical minimization techniques.

5 Illustration

In this section, we discuss a special case for which it is possible to obtain a closed-form expression for multivariate geometric expectiles. Moreover, we provide numerical illustrations for a number of different random vectors in order to highlight the impact of changing margins and dependence structures.

5.1. Analytic Solution for the Uniform Distribution

We consider the case of a bivariate uniform distribution and denote by $\mathbf{U} = (U_1, U_2)$ a random vector with density $\frac{1}{(b_1 - a_1)(b_2 - a_2)}$ where $b_j > a_j$ and $b_j, a_j \in \mathbb{R}$ for $j = 1, 2$. We first compute the expectation of the squared norm in terms of $\mathbf{c} = (c_1, c_2)$ as

$$\begin{aligned} g(c_1, c_2) &= \mathbb{E}[\|\mathbf{U} - \mathbf{c}\|_2^2] = \int_{a_1}^{b_1} \int_{a_2}^{b_2} \frac{(u_1 - c_1)^2 + (u_2 - c_2)^2}{(b_1 - a_1)(b_2 - a_2)} du_2 du_1 \\ &= \frac{1}{(b_1 - a_1)(b_2 - a_2)} \int_{a_1 - c_1}^{b_1 - c_1} \int_{a_2 - c_2}^{b_2 - c_2} u_1^2 + u_2^2 du_2 du_1 \\ &= \frac{(b_2 - a_2)((b_1 - c_1)^3 - (a_1 - c_1)^3) + (b_1 - a_1)((b_2 - c_2)^3 - (a_2 - c_2)^3)}{3(b_1 - a_1)(b_2 - a_2)}. \end{aligned}$$

Defining the real valued functions h_1 and h_2 as

$$\begin{aligned} h_1(x, y) &= \frac{1}{2} \left(y\sqrt{x^2 + y^2} + x^2 \log \left(y + \sqrt{x^2 + y^2} \right) \right) \quad \text{and} \\ h_2(x, y) &= \frac{1}{96} \left(-3x^4 + 20x^2 y \sqrt{x^2 + y^2} + y^3 \left(3y + 8\sqrt{x^2 + y^2} \right) + 12x^4 \log \left(y + \sqrt{x^2 + y^2} \right) \right), \end{aligned}$$

we further have that

$$\frac{d}{dx} h_2(x, y) = x h_1(x, y) \quad \text{and that} \quad \frac{d}{dy} h_1(x, y) = \sqrt{x^2 + y^2}.$$

Therefore,

$$\begin{aligned} \int_{a_1}^{b_1} x \int_{a_2}^{b_2} \sqrt{x^2 + y^2} dy dx &= \int_{a_1}^{b_1} x (h_1(x, b_2) - h_1(x, a_2)) dx \\ &= \int_{a_1}^{b_1} x h_1(x, b_2) dx - \int_{a_1}^{b_1} x h_1(x, a_2) dx \\ &= h_2(b_1, b_2) - h_2(a_1, b_2) - h_2(b_1, a_2) + h_2(a_1, a_2). \end{aligned}$$

This finally leads to

$$\begin{aligned} g_1(c_1, c_2) &= \mathbb{E}[\|\mathbf{U} - \mathbf{c}\|_2 (U_1 - c_1)] = \int_{a_1}^{b_1} \int_{a_2}^{b_2} \frac{(u_1 - c_1) \sqrt{(u_1 - c_1)^2 + (u_2 - c_2)^2}}{(b_1 - a_1)(b_2 - a_2)} du_2 du_1, \\ &= \frac{1}{(b_1 - a_1)(b_2 - a_2)} \int_{a_1 - c_1}^{b_1 - c_1} u_1 \int_{a_2 - c_2}^{b_2 - c_2} \sqrt{u_1^2 + u_2^2} du_2 du_1, \\ &= \frac{h_2(b_1 - c_1, b_2 - c_2) - h_2(a_1 - c_1, b_2 - c_2) - h_2(b_1 - c_1, a_2 - c_2) + h_2(a_1 - c_1, a_2 - c_2)}{(b_1 - a_1)(b_2 - a_2)}, \end{aligned}$$

Vector	Copula	$X_{i1} \sim$	$X_{i2} \sim$
$\mathbf{X}_1 = (X_{11}, X_{12})$	Independence	$\mathcal{N}(0, 1)$	$\mathcal{N}(0, 1)$
$\mathbf{X}_2 = (X_{21}, X_{22})$	Independence	$\mathcal{SN}(-1, 1, 2)$	t_4
$\mathbf{X}_3 = (X_{31}, X_{32})$	Gumbel, $\theta = 2$	$\mathcal{N}(0, 1)$	$\mathcal{N}(0, 1)$
$\mathbf{X}_4 = (X_{41}, X_{42})$	Gumbel, $\theta = 2$	$\mathcal{SN}(-1, 1, 2)$	t_4

Table 1: Specification of the random vectors $\mathbf{X}_1, \dots, \mathbf{X}_4$.

where we define g_2 analogously as $g_2(c_1, c_2) = \mathbb{E}[\|\mathbf{U} - \mathbf{c}\|_2 (U_2 - c_2)]$. Taking the preceding results together, we have for $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$ that

$$\begin{aligned} \phi(\mathbf{c}) &= \mathbb{E}[\Lambda_{\boldsymbol{\alpha}}(\mathbf{U} - \mathbf{c})] = \frac{1}{2} \mathbb{E}[\|\mathbf{U} - \mathbf{c}\|_2^2 + \alpha_1 \|\mathbf{U} - \mathbf{c}\|_2 (U_1 - c_1) + \alpha_2 \|\mathbf{U} - \mathbf{c}\|_2 (U_2 - c_2)] \\ &= \frac{g(c_1, c_2)}{2} + \frac{\alpha_1 g_1(c_1, c_2)}{2} + \frac{\alpha_2 g_2(c_1, c_2)}{2}. \end{aligned}$$

The geometric expectiles $e_{\boldsymbol{\alpha}}(\mathbf{U})$ are now found as

$$e_{\boldsymbol{\alpha}}(\mathbf{U}) = \operatorname{argmin}_{\mathbf{c} \in \mathbb{R}^2} \phi(\mathbf{c}).$$

This example highlights more than anything that finding a closed-form solution can be challenging even in the simplest of cases. In this sense, the numerical approximation introduced in Section 4.3 takes a more prominent role. Full-fledged examples utilizing this method can be found in the following sections.

5.2. Numerical Illustration

In this section we visualize geometric expectiles for selected bivariate random vectors. To this end, we define four random vectors $\mathbf{X}_1, \dots, \mathbf{X}_4$ with different margins and dependence structures; see Table 1. The dependence structure is formalized in terms of copulas, see, for example, Nelsen [2006] or Joe [2014] for textbook introductions. As a baseline for our comparison, $\mathbf{X}_1 = (X_{11}, X_{12})$ follows a bivariate normal distribution with independent standard normal margins. Considering \mathbf{X}_2 we keep the independence between the components, but we change the margins. X_{21} now follows a skew normal distribution, see Azzalini [1985], with parameters $(\xi, \omega, \alpha) = (-1, 1, 2)$ and X_{22} follows a Student t distribution with $\nu = 4$ degrees of freedom. In case of \mathbf{X}_3 we only change the dependence structure compared to \mathbf{X}_1 , that is X_{31} and X_{32} still follow a standard normal distribution each but the dependence structure is now given by a Gumbel copula with parameter $\theta = 2$. Finally \mathbf{X}_4 differs from \mathbf{X}_1 in terms of margins and dependence structure, where we employ the skew normal and Student t margins of \mathbf{X}_2 with the Gumbel dependence structure of \mathbf{X}_3 .

To illustrate the impact of different indices we consider two parameterizations for $\boldsymbol{\alpha}$. First, we choose $\boldsymbol{\alpha}$ according to $\boldsymbol{\alpha}_1(\varphi) = 0.98(\cos(\varphi), \sin(\varphi))$, $\varphi \in [0, 2\pi)$, which describes a circle

of radius 0.98. The magnitude 0.98 corresponds to a confidence level of 0.99 in the univariate case. Second, we choose α according to $\alpha_2(\varphi) = (0.98 \cos(\varphi), 0.90 \sin(\varphi))$, $\varphi \in [0, 2\pi)$, which describes an ellipse in B , where a magnitude of 0.90 corresponds to a confidence level of 0.95 in the univariate case. Both choices of indices are visualized in Figure 4. For further reference we indicate the resulting indices $\alpha_j(\varphi_k)$, $j \in \{1, 2\}$, for $\varphi_k = k2\pi/8$, $k \in \{0, \dots, 7\}$, by the respective value of k . As there are no closed-form solutions available to compute $e_{\alpha_j(\varphi)}(\mathbf{X}_\ell)$, $\ell \in \{1, 2, 3, 4\}$, we instead draw an iid sample of size 10,000 from the respective distribution of \mathbf{X}_ℓ and utilize the numerical procedure outlined in Section 4.3; i.e., we use Monte Carlo integration.

Figure 5 shows the resulting geometric expectiles and density contour lines for \mathbf{X}_1 (top left), \mathbf{X}_2 (top right), \mathbf{X}_3 (bottom left) and \mathbf{X}_4 (bottom right). The gray lines indicate the density contours of the underlying bivariate distribution function. To indicate the effects of different index choices the solid orange line represents the resulting geometric expectiles $e_{\alpha_1(\varphi)}(\mathbf{X}_\ell)$, $\ell \in \{1, \dots, 4\}$, for $\varphi \in [0, 2\pi)$. Likewise the solid green line indicates $e_{\alpha_2(\varphi)}(\mathbf{X}_\ell)$, $\ell \in \{1, \dots, 4\}$, for $\varphi \in [0, 2\pi)$. In concordance with Figure 4 we mark the resulting geometric expectiles $e_{\alpha_j(\varphi_k)}(\mathbf{X}_i)$ for indices $\alpha_j(\varphi_k)$ based on $\varphi_k = k2\pi/8$, $k \in \{0, \dots, 7\}$, by the respective value of k .

From Figure 5 it becomes apparent that geometric expectiles adapt to the underlying distribution. For the radially symmetric distribution of \mathbf{X}_1 (top left panel) the lines indicating $e_{\alpha_j(\varphi)}(\mathbf{X}_1)$ for all possible $\varphi \in [0, 2\pi)$ resemble the shape of the index $\alpha_j(\varphi)$. Furthermore we visually observe the symmetry established in Proposition 4.5. However, for skewed and heavier tailed margins (top right panel) the geometric expectiles adapt by bulging out. This also slightly changes the orientation in that, for example, $e_{\alpha_1(\varphi_2)}(\mathbf{X}_2)$ is not centered on the y -axis anymore. Introducing dependence between the components of \mathbf{X}_3 (bottom left panel) forces the geometric expectiles to deform. The deformation, compared to the top left panel, is, however, not by bulging out as in the top right panel, but rather by compressing and rotating. Finally when combining both effects in \mathbf{X}_4 (bottom right panel) we see that geometric expectiles widen and deform according to a superposition of the previously observed effects.

5.3. Comparing Geometric Value-at-Risk and Expectiles

In continuation of the numerical examples in Section 5.2 we now discuss differences between geometric VaR and geometric expectiles, as well as their univariate counterparts. For a fixed $\alpha_1 \in (0, 1)$ we therefore consider the corresponding index $\alpha = (2\alpha_1 - 1)(1, 0)$, where we make the necessary adjustment to the magnitude of the index discussed in Section 3.1. We then compute the univariate $\text{VaR}_{\alpha_1}(X_{11})$ and expectile $e_{\alpha_1}(X_{11})$ at level α_1 for the first component of $\mathbf{X}_1 = (X_{11}, X_{12})$, see Table 1, and also the geometric $\text{VaR}_\alpha(\mathbf{X}_1)$ and geometric expectile $e_\alpha(\mathbf{X}_1)$ based on α . Comparing the univariate risk measures to the first component of their multivariate counterparts in Figure 6, we see that the multivariate risk measures are more conservative, i.e., higher in absolute value. In fact, the geometric VaR provides the most conservative reserve estimates for a given level α_1 , while the univariate expectiles are the least conservative for the same level.

We further compare geometric VaR and geometric expectiles by computing the magnitude for a given direction that leads to equal values in each component of the resulting multivariate

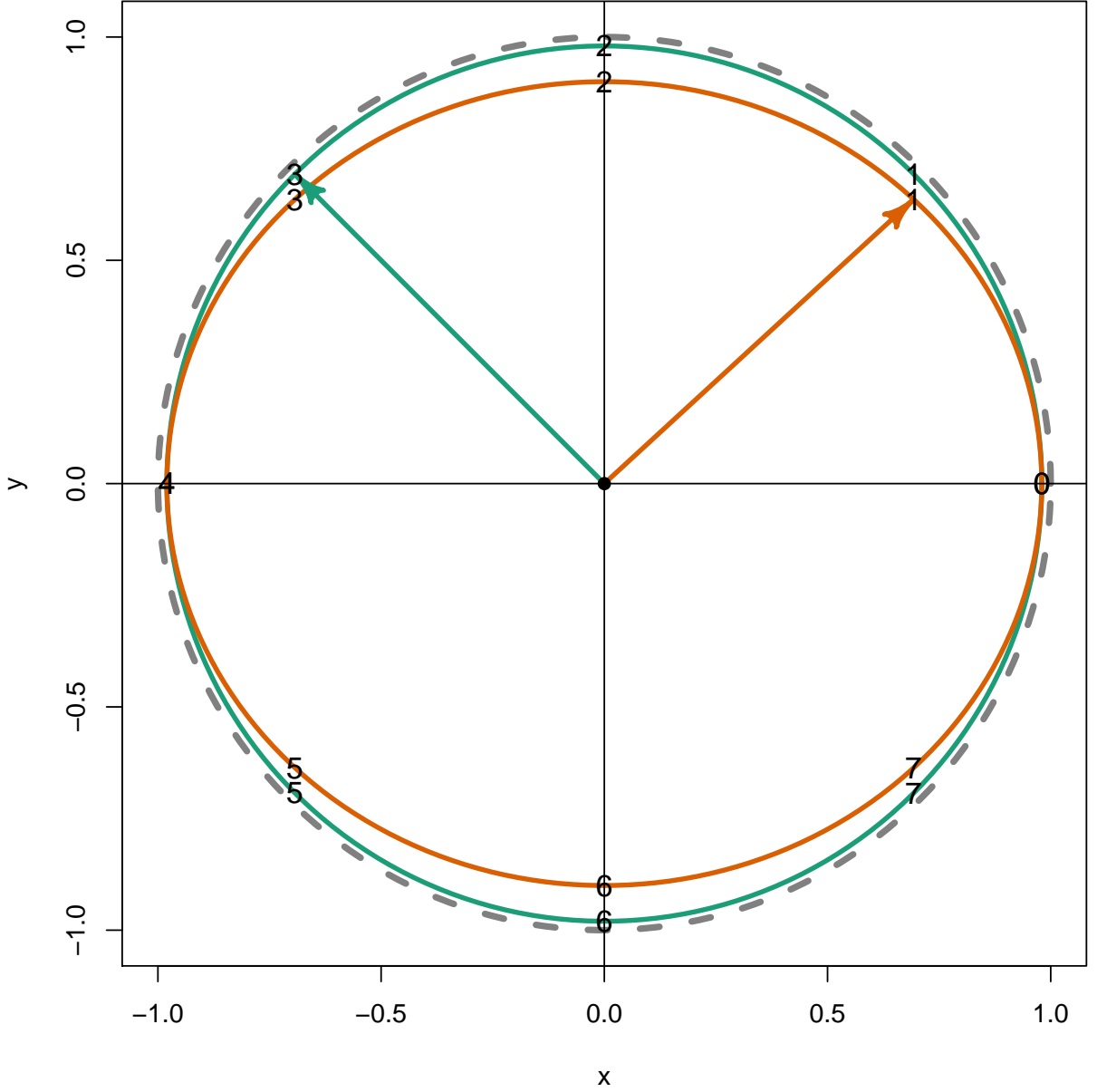


Figure 4: Bivariate indices $\alpha_1(\varphi) = 0.98(\cos(\varphi), \sin(\varphi))$ for $\varphi = 3\pi/4$ (green arrow) and $\alpha_2(\varphi) = (0.98 \cos(\varphi), 0.90 \sin(\varphi))$ (orange arrow) for $\varphi = \pi/4$. The solid lines indicate possible indices when φ varies in $[0, 2\pi)$. Numbers indicate $\alpha_j(\varphi_k)$, $j \in \{1, 2\}$, for $\varphi_k = k2\pi/8$ where $k \in \{0, \dots, 7\}$. The gray dashed line indicates the boundary of the open unit ball $\partial B = \{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\|_2 = 1\}$.

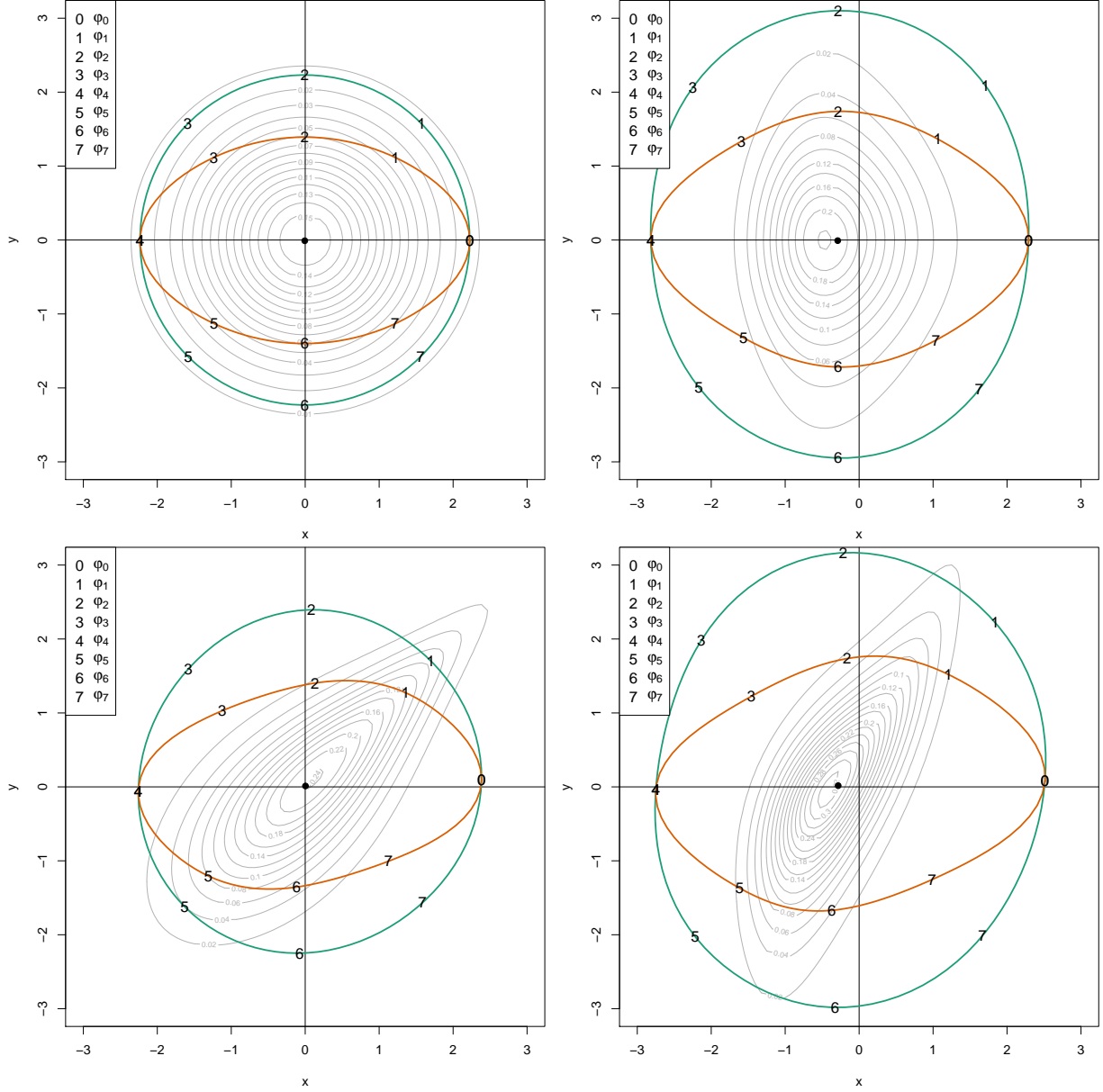


Figure 5: Geometric expectiles $e_{\alpha_j}(\mathbf{X}_\ell)$ for $\ell = 1$ (top left), $\ell = 2$ (top right), $\ell = 3$ (bottom left) and $\ell = 4$ (bottom right) and $j \in \{1, 2\}$. The density contour lines of \mathbf{X}_ℓ are given in gray. The solid green lines indicate $e_{\alpha_1(\varphi)}(\mathbf{X}_\ell)$, $\varphi \in [0, 2\pi)$, for $\alpha_1(\varphi) = 0.98(\cos(\varphi), \sin(\varphi))$. The solid orange lines indicate $e_{\alpha_2(\varphi)}(\mathbf{X}_\ell)$, $\varphi \in [0, 2\pi)$, for $\alpha_2(\varphi) = (0.98 \cos(\varphi), 0.90 \sin(\varphi))$. Numbers indicate $e_{\alpha_j(\varphi_k)}(\mathbf{X}_\ell)$ for $\varphi_k = k2\pi/8$ where $k \in \{0, \dots, 7\}$. The mean vector $(\mathbb{E}[X_{\ell,1}], \mathbb{E}[X_{\ell,2}])$ is represented by the black dot. Computations are based on 10,000 iid realizations of \mathbf{X}_ℓ for each $\ell \in \{1, \dots, 4\}$.

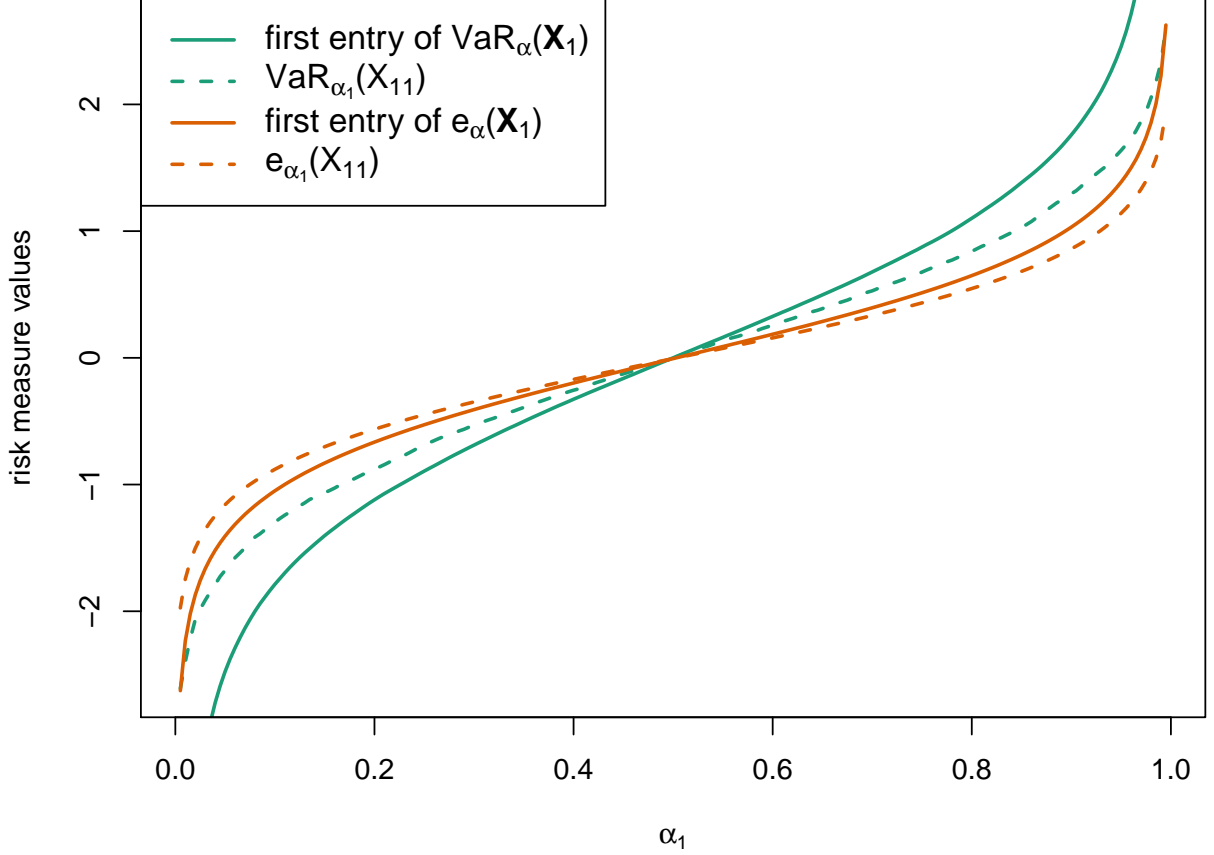


Figure 6: Solid lines indicate the first component of geometric VaR (green) and geometric expectiles (orange) based on $\alpha = (2\alpha_1 - 1)(1, 0)$ applied to $\mathbf{X}_1 = (X_{11}, X_{12})$, see Table 1. Dashed lines indicate univariate VaR (green) and expectiles (orange) at level α_1 applied to X_{11} .

risk measure. We therefore fix an element $\mathbf{u} \in \partial B = \{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\|_2 = 1\}$ and $\theta \in [0, 1)$ to obtain a starting index $\alpha = \theta\mathbf{u}$ for which we compute $e_\alpha(\mathbf{X}_1)$. For $m \in [0, 1)$ we then aim to find an optimal m^* that yields $\text{VaR}_{m\mathbf{u}}(\mathbf{X}_1) = e_\alpha(\mathbf{X}_1)$ in the least-square sense, that is

$$m^* = \underset{m \in [0, 1)}{\operatorname{argmin}} \sum_{i=1}^2 (\text{VaR}_{m\mathbf{u}}(\mathbf{X}_1)_i - e_\alpha(\mathbf{X}_1)_i)^2.$$

In Figure 7 we show the resulting plot for values of θ in $[0, 0.999]$ and $\mathbf{u} = (1, 1)/\sqrt{2}$. In line with Figure 6 we find that the associated magnitude m^* for the geometric VaR is lower than the corresponding magnitude θ for the geometric expectiles. That is to say that the geometric VaR is more conservative than geometric expectiles in this example.

5.4. Example Application

To demonstrate how geometric expectiles can be used in a practical scenario we consider a data generating process that generalizes the well-known compound Poisson model. By $\mathbf{E} =$

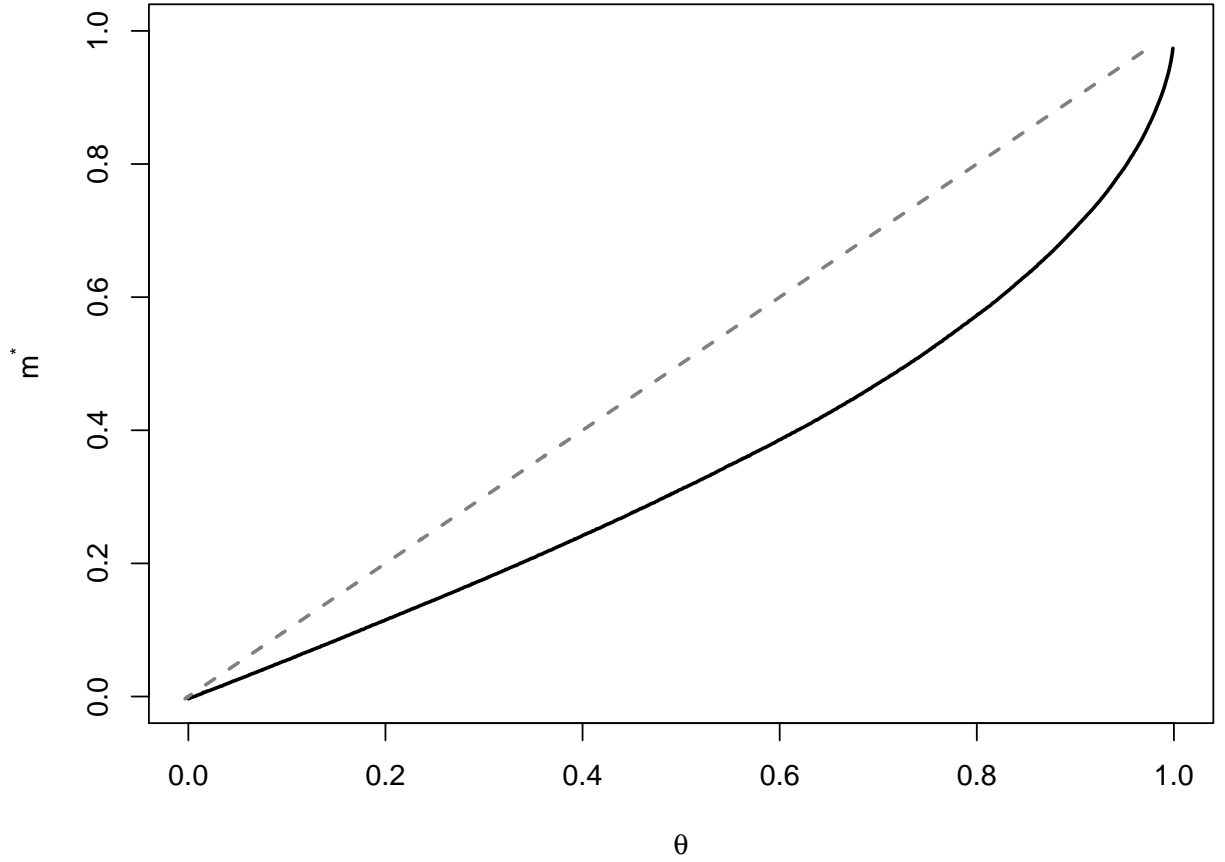


Figure 7: Pairs of magnitude (θ, m^*) such that $\text{VaR}_{m^*\mathbf{u}}(\mathbf{X}_1) = e_\alpha(\mathbf{X}_1)$ in the least-square sense. Here $\mathbf{u} = (1, 1)/\sqrt{2}$ and $\boldsymbol{\alpha} = \theta\mathbf{u}$. \mathbf{X}_1 is specified in Table 1. The gray-dashed line indicates the 45 degree line for comparison.

(E_1, E_2) we denote a random vector with exponentially distributed margins $E_\ell \sim \text{Exp}(\beta_\ell)$, $\ell = 1, 2$. The dependence structure between the components of \mathbf{E} is given by a Clayton copula C_θ with parameter $\theta > 0$. For a Poisson random variable $N \sim \text{Pois}(\lambda)$ our final random vector $\mathbf{X} = (X_1, X_2)$ is then given by

$$\mathbf{X} = \sum_{k=1}^N \mathbf{E}_k,$$

where \mathbf{E}_k is an independent (of N and \mathbf{E}_j for $j \neq k$) copy of \mathbf{E} . By construction we see that X_j , $j \in \{1, 2\}$, is a compound Poisson model with exponentially distributed severities. All in all the model captures the situation where a random number of risk occurs together, and the components of each incident are not independent. Our example is motivated by considering vehicle insurance that can¹ cover medical payments for the insured party as well as physical damages to the insured vehicle. From the point of the insurance company there will be a random number of accidents, where it is reasonable to assume a positive dependence between both components of the policy.

For our example we consider the parameters $\theta = 0.9$, $\beta_1 = 1/10$, $\beta_2 = 1/15$ and $\lambda = 1$. The computation of the geometric expectiles is now based on a simulated iid sample $(\mathbf{x}_i)_{i=1}^{100}$ of \mathbf{X} . The computation therefore utilizes the Monte Carlo estimator according to Corollary 4.1 and the discussion therein. Figure 8 shows the resulting geometric expectiles, where we again consider the previously introduced, see Figure 4 and Section 5.2, indices $\boldsymbol{\alpha}_1(\varphi) = 0.98(\cos(\varphi), \sin(\varphi))$ and $\boldsymbol{\alpha}_2(\varphi) = (0.98 \cos(\varphi), 0.90 \sin(\varphi))$. Given that in this example the margins are a.s. positive, we confine ourselves to directions in the first quadrant only, i.e., $\varphi \in [0, \pi/2]$. In this case numbers indicate the resulting geometric expectiles for indices $\boldsymbol{\alpha}_j(\varphi_k)$, $j \in \{1, 2\}$, where $\varphi_k = k\pi/14$, $k \in \{0, \dots, 7\}$.

Concerning the individual variables X_1 and X_2 the insurer can now reserve losses according to the indices $\boldsymbol{\alpha}_j(\varphi_0)$ or respectively $\boldsymbol{\alpha}_j(\varphi_7)$. Taking $j = 1$ corresponds to a traditional confidence level of 0.99 for both components, while $j = 2$ corresponds to a traditional confidence level of 0.99 for X_1 and 0.95 for X_2 . More importantly, by extending the univariate forecast model validation theory outlined, for example, in Gneiting [2011] or Nolde and Ziegel [2016] it might be possible to validate the proposed model against real data by backtesting. Using geometric expectiles the backtest would then validate the full joint distribution function of \mathbf{X} , and not just the individual marginal distributions of X_1 and X_2 .

6 Conclusion

In this paper we introduced geometric expectiles for multivariate distribution functions with finite marginal second moments. This proposed functional naturally generalizes univariate expectiles introduced in Newey and Powell [1987] to the multivariate case for any fixed dimension d . Instead of a single real number geometric expectiles are represented by a d -dimensional vector, an approach in line with other recently introduced multivariate risk

¹Coverage of medical costs depends on the respective jurisdiction. Vehicle insurance policies that cover medical and physical damage are common in the USA. On the other hand, there are, for example, no such products in the Québec province of Canada since medical costs are in this case taken over by the province.

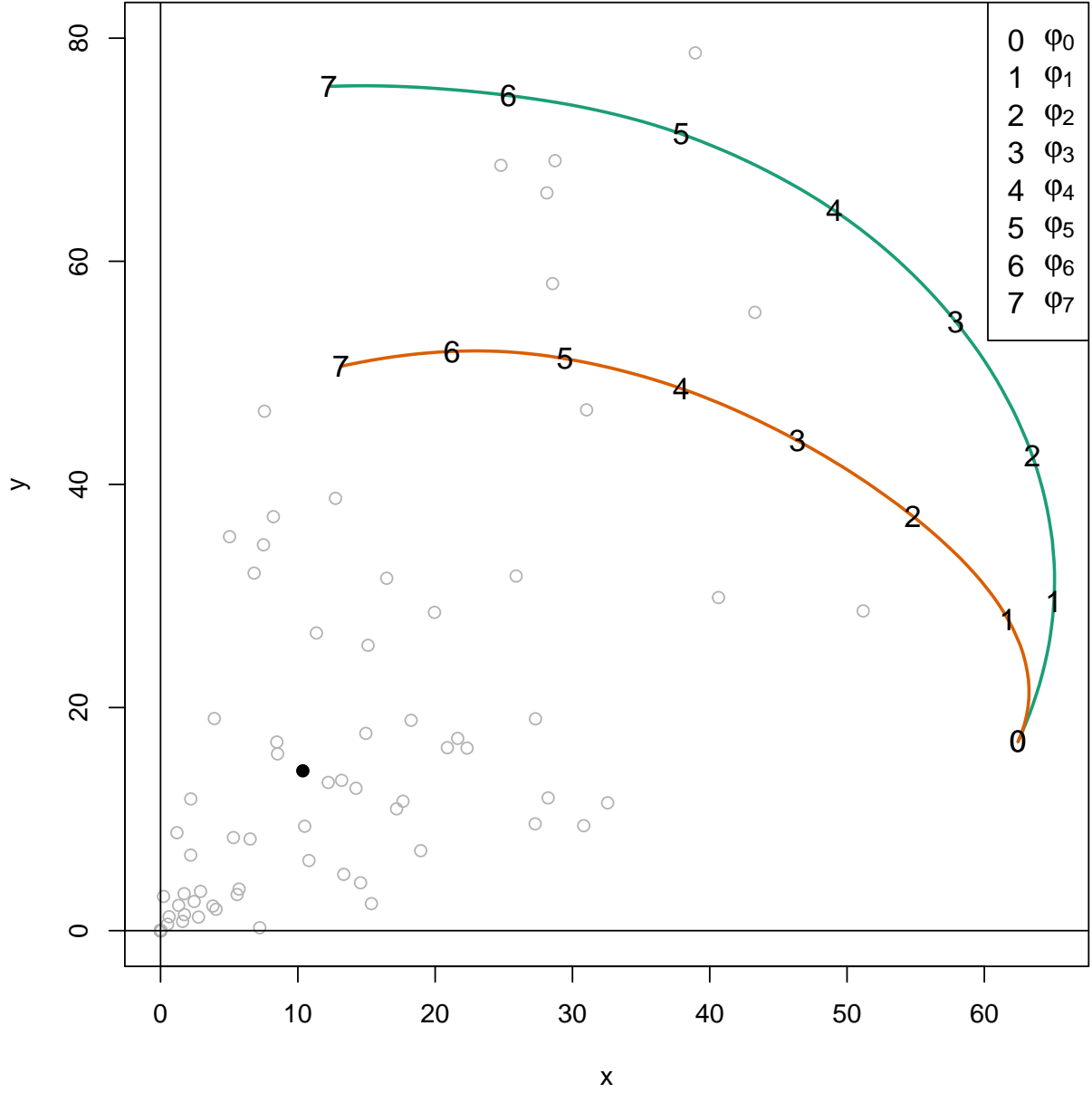


Figure 8: Geometric expectiles $e_{\alpha_1}(\mathbf{X})$ (green line) and $e_{\alpha_2}(\mathbf{X})$ (orange line) for the bivariate compound Poisson model. Indices are given as $\alpha_1(\varphi) = 0.98(\cos(\varphi), \sin(\varphi))$ and $\alpha_2(\varphi) = (0.98 \cos(\varphi), 0.90 \sin(\varphi))$, where $\varphi \in [0, \pi/2]$. Numbers indicate the resulting geometric expectiles for indices $\alpha_j(\varphi_k)$, $j \in \{1, 2\}$, where $\varphi_k = k\pi/14$, $k \in \{0, \dots, 7\}$. The black dot indicates the bivariate mean $(\mathbb{E}[X_1], \mathbb{E}[X_2])$. Computations are based on 100 iid realizations of \mathbf{X} marked by gray circles.

measures. Utilizing a framework comparable to the one introduced in Chaudhuri [1996] to generalize quantiles, the resulting geometric expectiles are indexed by an element of the open unit ball of \mathbb{R}^d .

Seen as a statistical functional geometric expectiles have a number of desirable properties. First, they are well-defined and unique for any multivariate distribution function with finite marginal second moments. Second, they are well-behaved under data transformations such as translating, re-scaling or re-ordering the data. Generalizing a re-ordering also multiplications with orthogonal matrices are well-behaved. Third, comparable to the univariate case, geometric expectiles are elicitable (in a multivariate sense). Comparable to the univariate case, this possibly provides one with a mechanism to rank competing multivariate forecasting procedures, or to backtest a multivariate model against real data.

Aside from population characteristics, we also studied properties and asymptotics of the corresponding finite sample version. Here we find that the sample version is a consistent estimator of the population characteristics. A Monte Carlo estimator of geometric expectiles is readily available when a closed-form solution is not.

In the presented examples, we utilized these simulation-based approximations to contrast geometric expectiles to the geometric quantiles introduced in Chaudhuri [1996] as well as univariate expectiles and quantiles. Our results indicate that geometric value-at-risk is more conservative than geometric expectiles for a given index.

In cases where the second moment condition on the margins is too restrictive it remains to be seen how tempering the margins interacts with geometric expectiles, providing a possible remedy.

Acknowledgements

This work was supported by NSERC under Grant RGPIN-5010-2015 and RGPIN-2015-05447. The authors would also like to thank an anonymous contributor for communicating a vital step in the proof of Theorem 4.2.

References

- B. Abdous and R. Theodorescu. Note on the spatial quantile of a random vector. *Statistics & Probability Letters*, 13(4):333–336, 1992.
- P. Artzner, F. Delbaen, J.M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999.
- A. Azzalini. A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, 12:171–178, 1985.
- A. Balbás, R. Balbás, and P. Jiménez-Guerra. Vector Risk Functions. *Mediterranean Journal of Mathematics*, 6:139–150, 2011.
- V. Barbu and T. Precupanu. *Convexity and Optimization in Banach Spaces*. Springer, 4th edition, 2012.

- F. Bellini and E. Di Bernardino. Risk management with expectiles. *The European Journal of Finance*, 23(6):487–506, 2017.
- I. Ben Tahar. Tail conditional expectation for vector-valued risks. *SFB 649 Discussion Papers*, 2006.
- B. Chakraborty. On affine equivariant multivariate quantiles. *Annals of the Institute of Statistical Mathematics*, 53(2):380–403, 2001.
- P. Chaudhuri. On a geometric notion of quantiles for multivariate data. *Journal of the American Statistical Association*, 91(434):862–872, 1996.
- U. Cherubini, E. Luciano, and W. Vecchiato. *Copula Methods in Finance*. John Wiley & Sons, 2004.
- H. Cossette, M. Mailhot, and E. Marceau. Tvar-based capital allocation for multivariate compound distributions with positive continuous claim amounts. *Insurance: Mathematics and Economics*, 50:247–256, 2012.
- H. Cossette, M. Mailhot, E. Marceau, and M. Mesfioui. Vector-valued tail value-at-risk and capital allocation. *Methodology and Computing in Applied Probability*, pages 1–22, 2015.
- A. Cousin and E. Di Bernardino. On multivariate extensions of value-at-risk. *Journal of Multivariate Analysis*, 119:32–46, 2013.
- A. Cousin and E. Di Bernardino. On multivariate extensions of conditional-tail-expectation. *Insurance: Mathematics and Economics*, 55:272–282, 2014.
- E. Di Bernardino, T. Laloë, V. Maume-Deschamps, and C. Prieur. Plug-in estimation of level sets in a non-compact setting with applications in multivariate risk theory. *ESAIM: Probability and Statistics*, 17:236–256, 2013.
- B. Efron. Regression percentiles using asymmetric squared error loss. *Statistica Sinica*, 1: 93–125, 1991.
- P. Embrechts and G. Puccetti. Bounds for functions of multivariate risks. *Journal of Multivariate Analysis*, 97(2):526–547, 2006.
- T. Fissler and J. Ziegel. Higher order elicibility and osband’s principle. *Annals of Statistics*, 44:1680–1707, 2016.
- C. Gauthier, A. Lehar, and M. Souissi. Macroprudential regulation and systemic capital requirements. Technical report, Bank of Canada Working Paper, 2010.
- T. Gneiting. Making and evaluating point forecasts. *Journal of the American Statistical Association*, 106(494):746–762, 2011.
- D. Guégan and B. Hassani. A mathematical resurgence of risk management: an extreme modeling of expert opinions. *Working paper*, 2011.

- M. Hallin and M. Paindaveine, D. Siman. Multivariate quantiles and multiple-output regression quantiles: From l1 optimization to halfspace depth. *The Annals of Statistics*, 38(2): 635–669, 2010.
- F. Hayashi. *Econometrics*. Princeton University Press, 2000.
- P.J. Huber and E.M. Ronchetti. *Robust Statistics*. Wiley, 2nd edition, 2009.
- H. Joe. *Dependence Modeling with Copulas*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis, 2014.
- E. Jouini, M. Meddeb, and N. Touzi. Vector-valued Coherent Risk Measures. *Finance and Stochastics*, 8:531–552, 2004.
- R. Koenker and J. R. Bassett. Regression quantiles. *Econometrica*, 46(1):33–50, 1978.
- E.L. Lehmann. *Theory of Point Estimation*. John Wiley & Sons, 1983.
- M. Mailhot, M. Morales, and H.F. Omid. Multivariate Database Natural Risk Statistics. *Submitted*, 2014.
- V. Maume-Deschamps, D. Rullière, and K. Said. Multivariate extensions of expectiles risk measures. *Dependence Modeling*, 5:20–44, 2017.
- A. McNeil and A. Smith. Multivariate stress scenarios and solvency. *Insurance: Mathematics and Economics*, 50:299–308, 2012.
- A. McNeil, R. Frey, and P. Embrechts. *Quantitative Risk Management: Concepts, Techniques, and Tools*. Princeton University Press, 2nd edition, 2015.
- K. Mosler. Depth statistics. In Fried C. and Kuhnt and R., editors, *Becker*, pages 17–34. Robustness and Complex Data Structures, Festschrift in Honour of Ursula Gather. Springer, Berlin, 2013.
- R.B. Nelsen. *An introduction to copulas*. Springer Verlag, 2006.
- W. Newey and J. Powell. Asymmetric least squares estimation and testing. *Econometrica*, 55(4):819–847, 1987.
- N. Nolde and J. Ziegel. Elicitability and backtesting. *Working paper*, 2016. URL <https://arxiv.org/abs/1608.05498>.
- Office of the Superintendent of Financial Institutions. Memorandum for the appointed actuary’s report on property and casualty insurance business. 2014. URL www.osfi-bsif.gc.ca.
- Kent Osband. *Providing incentives for better cost forecasting*. PhD thesis, University of California, Berkeley, 1985.
- OSFI. Key principles for the future direction of the canadian regulatory capital framework for property & casualty insurance. 2010. URL www.osfi-bsif.gc.ca.

- A. J. Patton. Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160:246–256, 2011.
- W. Rudin. *Principles of Mathematical Analysis*. McGraw-Hill, 3rd edition, 1976.
- R. Serfling. Quantile functions for multivariate analysis: approaches and applications. *Statistica Neerlandica*, 56(2):214–232, 2002.
- R. Torres, R. E. Lillo, and H. Laniado. A directional multivariate value at risk. *Insurance: Mathematics and Economics*, 65:111–123, 2015.
- H. Tuy. *Convex Analysis and Global Optimization*. Springer, 2nd edition, 2016.
- J. Ziegel. Coherence and elicibility. *Mathematical Finance*, 26(4):901–918, 2014.